

Parton distributions confront LHC Run II data: a quantitative appraisal

Amedeo Chiefa ^a, **Mark N. Costantini** ^b, **Juan Cruz-Martinez** ^c,
Emanuele R. Nocera ^{d,e}, **Tanjona R. Rabemananjara** ^{f,g}, **Juan Rojo** ^{f,g},
Tanishq Sharma ^{d,e,f,g}, **Roy Stegeman** ^a and **Maria Ubiali** ^b

^a*The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland, U.K.*

^b*DAMTP, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.*

^c*CERN, Theoretical Physics Department, CH-1211 Geneva 23, Switzerland*

^d*Dipartimento di Fisica, Università degli Studi di Torino, via Pietro Giuria 1, 10125 Torino, Italy*

^e*INFN, Sezione di Torino, via Pietro Giuria 1, 10125 Torino, Italy*

^f*Department of Physics and Astronomy, Vrije Universiteit, NL-1081 HV Amsterdam, The Netherlands*

^g*Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

*E-mail: amedeo.chiefa@ed.ac.uk, mnc33@cam.ac.uk,
juan.cruz.martinez@cern.ch, emanueleroberto.nocera@unito.it,
rrabeman@nikhef.nl, j.rojo@vu.nl, tanishq.sharma@edu.unito.it,
r.stegeman@ed.ac.uk, M.Ubiali@damtp.cam.ac.uk*

ABSTRACT: We present a systematic comparison of theoretical predictions and various high-precision experimental measurements, specifically of differential cross sections performed by the LHC run II for Drell-Yan gauge boson, top-quark pair, single-inclusive jet and di-jet production, and by HERA for single-inclusive jet and di-jet production. Theoretical predictions are computed at next-to-next-to-leading order (NNLO) accuracy in perturbative Quantum Chromodynamics. The most widely employed sets of Parton Distribution Functions (PDFs) are used, and PDF, strong coupling, and missing higher order uncertainties are taken into account. We quantitatively assess the predictive power of each PDF set and the contribution of the different sources of experimental and theoretical uncertainty to the agreement between data and predictions. We show that control over all of these aspects is crucial to precision physics studies, such as the determination of Standard Model parameters at the LHC.

KEYWORDS: Parton Distributions, Specific QCD Phenomenology

ARXIV EPRINT: [2501.10359](https://arxiv.org/abs/2501.10359)

Contents

1	Introduction	1
2	Experimental data and theoretical predictions	3
2.1	Drell-Yan weak boson production at the LHC	5
2.2	Top quark pair production at the LHC	6
2.3	Jet production at the LHC	7
2.4	Jet production at HERA	8
3	Methodological framework	9
3.1	Figure of merit	9
3.2	Stability of the experimental covariance matrix	12
3.3	PDF sets	14
4	Data-theory comparison appraisal	17
4.1	Drell-Yan weak boson production measurements at 13 TeV	19
4.2	The ATLAS 8 TeV inclusive Z boson production measurement	22
4.3	Top-quark pair production measurements	25
4.4	Single-inclusive jet and di-jet production measurements at the LHC	29
4.5	Single-inclusive jet and di-jet production measurements at HERA	32
4.6	Combined interpretation	33
5	Summary and outlook	36
A	Impact of the experimental covariance matrix regularisation	39
B	Additional results	42

1 Introduction

The remarkable progress witnessed by the determination of the parton distribution functions (PDFs) of the proton [1–4] in recent years has been driven by three factors: the extension of the input dataset, in particular thanks to high-precision Large Hadron Collider (LHC) measurements; the improvement of the accuracy of theoretical computations, now reaching the approximate next-to-next-to-next-to-leading order (aN³LO) in the strong coupling; and the investigation of several methodological aspects, specifically with respect to the quantification of PDF uncertainties. Some groups [5–8] provide regular updates of their PDF determinations based on a broad input dataset, while other groups focus on the interpretation of restricted subsets of data [9–11]. All of these PDF sets differ, sometimes by amounts that are larger than their nominal uncertainties, which can have very disparate size in different PDF sets. In order to understand the origin of these differences, several benchmark studies have been performed over the years [12–27]. In particular, it was shown [26] that PDFs determined from an identical

set of experimental data and theoretical predictions with three different methodological frameworks [6, 7, 28] displayed similar central values but somewhat different uncertainties. These differences should likely be ascribed to the different methodological framework.

These differences challenge the theoretical interpretation of the outcome of PDF determinations in terms of the underlying proton structure, as highlighted by the ongoing discussion concerning the possible existence of intrinsic charm quarks in the proton [29–31]. Furthermore, they degrade the physics reach of core LHC analyses sensitive to PDFs, concerning both the measurements of fundamental Standard Model (SM) parameters — like the strong coupling $\alpha_s(m_Z)$, the W -boson mass m_W , and the effective leptonic mixing angle $\sin^2 \theta_{\text{eff}}^\ell$ — and direct (resonance-like) and indirect (effective-field-theory (EFT)-like) searches for physics beyond the SM. The first aspect is illustrated, for instance, by the large PDF dependence of the high-mass forward-backward asymmetry in Drell-Yan gauge boson production [32–34]. The second aspect is illustrated by the interplay between PDFs and possible EFT contamination in high- p_T top-quark pair and Drell-Yan cross section measurements [35–41].

Recent LHC analyses have highlighted this far-from-ideal state of affairs. Here are three examples. First, the ATLAS determination of the strong coupling $\alpha_s(m_Z)$ from neutral-current Drell-Yan differential measurements in the transverse momentum of the Z boson [42]. This is the most precise α_s determination ever performed from a single experiment, with a quoted uncertainty $\delta = 9 \cdot 10^{-4}$. Of this value, the uncertainty due to the PDF is estimated to be the dominant component, $\delta_{\text{pdf}} = 5 \cdot 10^{-4}$, using the MSHT20 aN³LO fit [43]. However, if the PDF uncertainty is defined as the difference between central predictions obtained with the CT18A [6] and NNPDF4.0 [8] PDF sets, one gets an uncertainty which is four times larger, $\delta_{\text{pdf}} = 2 \times 10^{-3}$. Second, the CMS measurement of the effective leptonic mixing angle $\sin^2 \theta_{\text{eff}}^\ell$ [44]. In this case, the PDF uncertainty is estimated to be $\delta_{\text{pdf}} = 0.14\%$ using the CT18Z PDF set [6], while the spread of the central results obtained with CT18 [6] and MSHT [7] (before profiling) is around a factor of five larger, $\delta_{\text{pdf}} = 0.7\%$. Third, the updated ATLAS measurement of the W mass at 7 TeV [45]. In this case, the PDF uncertainty is estimated to be $\delta_{\text{pdf}} = 7.7$ (14.6) MeV in the lepton transverse momentum p_T^ℓ (transverse mass m_T) channel, to be compared with the spread between NNPDF4.0 and MSHT20 which gives twice that estimate, $\delta_{\text{pdf}} = 17$ (21) MeV. Similar considerations apply to the precise m_W measurement performed by the CMS collaboration [46]. Each of these analyses selects a different baseline PDF set, with other possible choices of PDFs yielding a central value potentially outside the quoted PDF uncertainties. Finally — and crucially — these analyses do not satisfactorily consider the back-reaction of the precision measurement under scrutiny on all other datasets entering a PDF fit [47], which is especially relevant when profiling/reweighting techniques for *in situ* calibration are used.

These considerations highlight the importance of understanding the origin of the differences observed when computing theoretical predictions with different PDF sets. Complementing existing benchmark studies that tackle this question, here we investigate whether existing PDF sets can be discriminated according to their predictive power of high-precision measurements not included in their determination. We will specifically consider cross sections measured by the LHC run II for Drell-Yan gauge boson, top-quark pair, single-inclusive jet and di-jet production, and by HERA for single-inclusive jet and di-jet production. We

will compare these experimental data to theoretical predictions computed at next-to-next-to-leading order (NNLO) accuracy in perturbative Quantum Chromodynamics (QCD) and quantitatively assess their mutual agreement. We will take into account all sources of theoretical uncertainty in this assessment, namely PDF, α_s , and missing higher order (MHO) uncertainties. We will study the dependence of this assessment on the input PDF set. This exercise is an extension of the future test introduced in [48].

The outline of this paper is as follows. In section 2 we present the considered LHC and HERA measurements and the computation of the corresponding theoretical predictions. In section 3 we describe how we quantitatively assess the agreement between experimental data and theoretical predictions, and in particular how we account for PDF, α_s , and MHO uncertainties in this assessment. In section 4 we present a selection of representative results for each class of measurements, highlighting the relative contribution of the various sources of theoretical uncertainty in the description of the data, and commenting on features that are common to or different from various PDF sets. We summarise our findings in section 5. Two appendices complete the paper. Appendix A quantifies the impact of regularizing ill-conditioned experimental covariance matrices in the assessment of the data-theory comparison. Appendix B collects the complete set of results not shown in section 4.

2 Experimental data and theoretical predictions

In this section, we present the experimental data considered in this work and the details of the corresponding theoretical computations. The data has been selected according to the following criteria.

- We consider datasets for scattering processes that provide information on PDFs of different partons (quarks, antiquarks, gluon) in a broad kinematic region of x and Q^2 . For a given process, we select the dataset based on the largest integrated luminosity available.
- We avoid datasets that are already included in PDF determinations used to compute theoretical predictions, to avoid double-counting. The only exception is the recent re-analysis of the Z data at a centre-of-mass energy of 8 TeV by ATLAS [42].
- We consider datasets for which the corresponding theoretical predictions can be computed at NNLO in perturbative QCD using event generators interfaced to fast interpolation grids. This avoids reliance on K -factors and allows one to readily evaluate predictions upon changes of input PDF set and factorisation and renormalisation scales.
- We only consider datasets for which the corresponding experimental information is publicly available, in particular through the HEPDATA repository [49].

Taking into account these requirements, the ATLAS, CMS, LHCb, H1, and ZEUS datasets that are considered in this study are summarised in table 1, classified by process type. For each dataset we indicate the experiment, the final-state channel, the measured differential distribution(s), the centre-of-mass energy, the integrated luminosity, the number of data points (after kinematic cuts), and the corresponding publication reference. For the

Process	Experiment	Final State	Observable	\sqrt{s} (TeV)	\mathcal{L} (fb $^{-1}$)	n_{dat}	Ref.	
LHC W, Z	ATLAS	Z p_T spectrum	$(\frac{1}{\sigma}) \frac{d\sigma}{dp_T^\ell}$	13	36.1	38	[50]	
	CMS	W incl. prod.	$\frac{d\sigma}{d \eta }$	13	35.9	36	[51]	
	LHCb	Z incl. forward prod.	$\frac{d\sigma}{dy^Z}$	13	5.1	18	[52]	
	ATLAS	Z incl. prod.	$\frac{d\sigma}{d y }$	8	20.2	7	[42]	
top-pair	ATLAS	all-hadronic	$(\frac{1}{\sigma}) \frac{d\sigma}{dm_{t\bar{t}}}$	13	36.1	9	[53]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{d y_{t\bar{t}} }$	13	36.1	12	[53]	
			$(\frac{1}{\sigma}) \frac{d^2\sigma}{d y_{t\bar{t}} dm_{t\bar{t}}}$	13	36.1	11	[53]	
	ATLAS	ℓ +jets	$(\frac{1}{\sigma}) \frac{d\sigma}{dm_{t\bar{t}}}$	13	36.1	9	[54]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{dp_T^\ell}$	13	36.1	8	[54]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{d y_t }$	13	36.1	5	[54]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{d y_{t\bar{t}} }$	13	36.1	7	[54]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{dm_{t\bar{t}}}$	13	137	15	[55]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{dp_T^\ell}$	13	137	16	[55]	
			$(\frac{1}{\sigma}) \frac{d\sigma}{d y_{t\bar{t}} }$	13	137	10	[55]	
CMS	ℓ +jets	$(\frac{1}{\sigma}) \frac{d\sigma}{d y_t }$	13	137	11	[55]		
		$(\frac{1}{\sigma}) \frac{d^2\sigma}{d y_{t\bar{t}} dm_{t\bar{t}}}$	13	137	35	[55]		
		ATLAS	incl. jet $R = 0.4$	$\frac{d^2\sigma}{dp_T d y }$	13	3.2	177	[56]
		CMS	incl. jets $R = 0.4$ (0.7)	$\frac{d^2\sigma}{dp_T d y }$	13	36.3 (33.5)	78	[57]
ATLAS	di-jets $R = 0.4$	$\frac{d^2\sigma}{dm_{jj} d y^* }$	13	3.2	136	[56]		
HERA jets	H1	incl. jet (low Q^2)	$\frac{d^2\sigma}{dQ^2 dp_T}$	0.319	0.29	37	[58]	
	H1	incl. jet (high Q^2)	$\frac{d^2\sigma}{dQ^2 dp_T}$	0.319	0.351	24	[59]	
	ZEUS	incl. jet	$\frac{d^2\sigma}{dQ^2 dE_T}$	0.300	0.038	30	[60]	
	ZEUS	incl. jet	$\frac{d^2\sigma}{dQ^2 dE_T}$	0.319	0.082	30	[61]	
	H1	di-jets (low Q^2)	$\frac{d^2\sigma}{dQ^2 d(p_T)}$	0.319	0.29	37	[58]	
	H1	di-jets (high Q^2)	$\frac{d^2\sigma}{dQ^2 d(p_T)}$	0.319	0.351	24	[59]	
	ZEUS	di-jets	$\frac{d^2\sigma}{dQ^2 d(E_T)}$	0.319	0.374	22	[62]	

Table 1. The ATLAS, CMS, LHCb, H1, and ZEUS datasets considered in this work, classified by process type. For each dataset we indicate the experiment, the final-state channel, the measured differential distribution(s), the centre-of-mass energy \sqrt{s} , the integrated luminosity \mathcal{L} , the number of data points n_{dat} (after kinematic cuts), and the corresponding publication reference. For the CMS single-inclusive jet production and for the ATLAS and CMS top-quark pair production datasets, we list the separate distributions provided by the corresponding analyses.

ATLAS and CMS top-quark pair production and for the CMS single-inclusive jet production datasets, we list all the separate distributions provided by the corresponding analyses. In this work, we select a subset of these distributions, which we deem most representative as explained in section 3.2. In the following, we discuss the main features of these datasets and describe the associated theoretical calculations.

2.1 Drell-Yan weak boson production at the LHC

Neutral- and charged-current Drell-Yan production is used to probe quark-flavour PDF separation, through rapidity distributions in the central (ATLAS and CMS) and forward (LHCb) regions [63, 64], and the gluon PDF, through transverse momentum distributions [65]. In the former case, the leading partonic channel is initiated by quarks and antiquarks; in the latter case, a non-zero p_T distribution arises from the $gq(\bar{q})$ partonic initial state followed by a hard $g \rightarrow q\bar{q}$ splitting. Here we consider three LHC Run II representative measurements for each of these categories: one by ATLAS [50], one by CMS [51], and one by LHCb [66]. All these measurements correspond to a centre-of-mass energy of 13 TeV. We also consider the recent re-analysis of the inclusive Z boson production measurement at a centre-of-mass energy of 8 TeV by ATLAS extrapolated to the full leptonic phase space [42].

The ATLAS Run II measurement [50] corresponds to an integrated luminosity of 36.1 fb^{-1} . It consists of the Z -boson production cross section, reconstructed from the combination of events resulting from electron and muon decays, differential in the transverse momentum of the dilepton pair $p_T^{\ell\ell}$. The measurement is performed in a fiducial phase space, defined by the lepton transverse momentum $p_T^\ell > 27 \text{ GeV}$, the absolute lepton pseudorapidity $|\eta_\ell| < 2.5$, and the dilepton invariant mass $66 < m_{\ell\ell} < 116 \text{ GeV}$. Cross sections are provided for both the absolute distribution and the distribution normalised to the fiducial cross section. The full breakdown of correlated systematic uncertainties is available and taken into account.

The CMS Run II measurement [51] corresponds to an integrated luminosity of 35.9 fb^{-1} . It consists of the W^\pm boson production cross section, reconstructed from the combination of events resulting from electron and muon decays. This measurement is presented as a double-differential distribution in the absolute lepton rapidity $|\eta|$, with $|\eta| < 2.4$, and in the lepton transverse momentum p_T^ℓ , with $26 < p_T^\ell < 56 \text{ GeV}$. It is available for each W polarisation state and averaged over polarisations. For each boson, 18 equally large bins in $|\eta|$ and a single bin in p_T^ℓ are provided. The full breakdown of correlated systematic uncertainties is available and taken into account.

The LHCb Run II measurement [52] corresponds to an integrated luminosity of 5.1 fb^{-1} . It consists of the Z -boson production cross section, reconstructed only from muon decays, in the fiducial phase space defined by the muon transverse momentum $p_T^\mu > 20 \text{ GeV}$, the dimuon invariant mass $60 < m_{\mu\mu} < 120 \text{ GeV}$, and the muon rapidity $2.0 < \eta_\mu < 4.5$. The presented cross section is differential in the rapidity of the Z boson y^Z . The full breakdown of correlated systematic uncertainties is available and taken into account.

Finally, we consider the recent ATLAS measurement of Z boson production based on the 2012 dataset at a centre-of-mass energy of 8 TeV, which corresponds to an integrated luminosity of 20.2 fb^{-1} [42]. The measurement is extrapolated to the full phase space of the decay electrons and muons in the dilepton rapidity range $|y| < 3.6$, and covers the Z pole invariant mass region, $80 \leq m_{\ell\ell} \leq 100 \text{ GeV}$. We specifically consider the cross section differential in $|y|$. The dependence on the transverse momentum of the dilepton pair is integrated over. The precision of this measurement, excluding the luminosity uncertainty, ranges from 0.2%, for $|y| \leq 2.0$, to 0.9% at more forward rapidities. This measurement is based on a re-analysis of events that were previously used in another two measurements [67, 68] from which double- and triple-differential distributions in the fiducial region for the final-state

leptons were reconstructed. The distributions are differential, respectively, in the invariant mass $m_{\ell\ell}$ and absolute rapidity $|y|$ of the dilepton pair, and in $m_{\ell\ell}$, $|y|$, and the cosine of the Collins-Soper angle, $\cos\theta^*$. The covered invariant mass region extends below, across and above the Z peak. The double differential measurement was included in the MSHT20 [7] and NNPDF4.0 [8] PDF fits. For this reason, the new measurement [42] does not fulfil the second selection criterion established at the beginning of this section. We make an exception for this measurement because, first, it exhibits a significant PDF dependence, and, second, it underlies the most precise determination of the strong coupling ever performed at a hadron collider, in which PDF uncertainties are the leading uncertainties. In section 4.2 we will discuss the interplay between the original [67, 68] and new [42] measurements.

For all these measurements, theoretical predictions accurate to NNLO QCD are computed in the form of PINEAPPL interpolation grids [69] with NNLOJET [70, 71]. The computation incorporates in particular the NNLO QCD corrections to the transverse momentum distributions of the Z boson from refs. [72, 73]. The central renormalisation and factorisation scales are set to

$$\mu_F = \mu_R = \sqrt{m_{\ell\ell}^2 + (p_T^{\ell\ell})^2}, \quad \mu_F = \mu_R = M_V, \quad (2.1)$$

respectively for the Z transverse momentum distribution and the gauge boson rapidity distributions (with M_V the gauge boson mass, Z or W). In the former case, we also apply a kinematic cut $p_T^{\ell\ell} > 30$ GeV to remove the region where resummation corrections, not accounted for by our fixed-order computation, may be relevant [65, 74]. Electroweak, QED, and photon-induced corrections, though known, are not considered here.

2.2 Top quark pair production at the LHC

Top-quark pair production at the LHC, which is initiated by gluon-gluon scattering, primarily probes the gluon PDF at large x [37, 75–81]. In addition to their PDF sensitivity, top-quark pair cross sections also constrain the top-quark mass m_t and the strong coupling $\alpha_s(m_Z)$ [82, 83]. Here we consider the ATLAS measurement reconstructed from the all-hadronic [53] and lepton+jets [54] final states, and the CMS measurement reconstructed from the lepton+jets final state [55]. We consider only parton-level measurements presented in terms of the kinematic variables of the final-state top quarks. The reason being that theoretical computations accurate to NNLO QCD for particle-level measurements [84] are not available in a numerical format suitable for this analysis. All these measurements were taken during the LHC Run II, at a centre-of-mass energy of 13 TeV.

The ATLAS measurements correspond to an integrated luminosity of 36 fb^{-1} . Cross sections are provided, absolute and normalised to the total cross section, as single- and double-differential distributions in various kinematic variables. For the sake of this work, we consider a subset of them, either the normalised or the absolute differential distributions. The choice depends on the stability of the experimental covariance matrix, as we will explain in section 3.2. For the all-hadronic measurement, we choose the single-differential absolute (normalised) distributions in the invariant mass (absolute rapidity) of the top-quark pair, $m_{t\bar{t}}$ ($y_{t\bar{t}}$), and the double-differential absolute distribution in these two variables. For the lepton+jets measurement, we choose the single-differential normalised distributions in the invariant mass

of the top quark pair, $m_{t\bar{t}}$, in the transverse momentum of the top quark, p_T^t , in the absolute rapidity of the top-quark pair, $y_{t\bar{t}}$, and in the absolute rapidity of the top-quark, y_t .

The CMS measurement corresponds to an integrated luminosity of 137 fb^{-1} , that is, all the events recorded during the LHC Run II. Absolute and normalised cross sections are provided as single- and double-differential distributions in various kinematic variables. We select a subset of them, specifically the single-differential normalised distributions in $m_{t\bar{t}}$, $|y_{t\bar{t}}|$, $|y_t|$, and p_T^t , and the double-differential normalised distribution in $(m_{t\bar{t}}, |y_{t\bar{t}}|)$.

Theoretical predictions accurate to NNLO QCD are computed using MATRIX [85], which has been interfaced to PINEAPPL [69]. The central factorisation and renormalisation scales are set to

$$\mu_R = \mu_F = H_T/2 = \sqrt{m_t^2 + (p_T^t)^2}/2, \tag{2.2}$$

where m_t and p_T^t are the mass and the transverse momentum of the top quark. This choice follows the recommendation of [86]. A value of $m_t^{\text{pole}} = 172.5\text{ GeV}$ has been used for the top-quark pole mass, consistently with the latest PDG average [87]. These computations have been benchmarked, when possible, against FASTNLO tables [88] generated with the code presented in [89]. Electroweak, QED, and photon-induced cross sections are not included.

2.3 Jet production at the LHC

Collider jet production at the LHC is a traditional probe of the gluon PDF in the large- x region [90–93], though it provides also information on the large- x quark PDFs. Here we consider the ATLAS measurement of single-inclusive jet and di-jet production [56], and the CMS measurement of single-inclusive jet production [57]. Both of them were taken during the LHC Run II, at a centre-of-mass energy of 13 TeV.

The ATLAS measurements correspond to an integrated luminosity of 3.2 fb^{-1} . Whereas this amounts to only a small fraction of the events recorded during Run II, no other unfolded measurements of single-inclusive jet or di-jet production based on a larger sample have been presented by ATLAS to date. The single-inclusive jet measurement is presented as a set of double differential cross sections in the jet transverse momentum p_T , with $100\text{ GeV} \leq p_T \leq 3.5\text{ TeV}$, and the jet absolute rapidity $|y|$, with $|y| < 3.0$. The di-jet measurement is presented as a set of double differential cross sections in the di-jet invariant mass m_{jj} , with $300\text{ GeV} \leq m_{jj} \leq 9\text{ TeV}$, and the half absolute rapidity separation between the two leading jets, $|y^*|$, with $|y^*| < 3.0$. Single-inclusive jets and di-jets are reconstructed by means of the anti- k_T clustering algorithm [94] for a jet radius of $R = 0.4$. The full breakdown of correlated systematic uncertainties is available, separately for the single-inclusive jet and di-jet measurements, and taken into account.

The CMS measurement corresponds to an integrated luminosity of 36.3 (33.5) fb^{-1} and a jet radius of $R = 0.4$ ($R = 0.7$). We consider the measurement with $R = 0.4$. Cross sections are double differential in the jet transverse momentum p_T , with $97 \leq p_T \leq 3.1\text{ TeV}$, and in the jet absolute rapidity $|y|$, with $|y| < 2.0$. The experimental covariance matrix of the measurement is provided and taken into account.

For all the aforementioned measurements, theoretical predictions, accurate to NNLO QCD in the leading color approximation, were computed with the NNLOJET code [95]. The

central factorisation and renormalisation scales where chosen as

$$\mu_F = \mu_R = p_T, \quad \mu_F = \mu_R = m_{jj}, \quad (2.3)$$

respectively for single-inclusive jets and di-jets. These predictions were released [96] as interpolation grids in the APPLFAST format through the PLOUGHSHARE website [97]. For the sake of this work, these grids have been converted to the PINEAPPL format [69]. We do not account for NLO electroweak corrections or photon-initiated contributions, neither for single-inclusive jets nor for di-jets. Monte Carlo uncertainties due to the generation of a finite number of events are generally smaller than MHO and α_s uncertainties, and are thus ignored.

2.4 Jet production at HERA

Jet production in DIS can probe the gluon PDF at large x as well. This process was measured at HERA by the H1 and ZEUS experiments and demonstrated to constrain the gluon PDF and the strong coupling [98–100] in comparison to fits based on inclusive DIS measurements only. Here we consider four H1 measurements [58, 59, 101] and three ZEUS measurements [60–62] for single-inclusive jet and di-jet cross sections, as indicated in table 1.

The H1 measurements correspond to the HERA-II data-taking period with a centre-of-mass energy of 319 GeV. Two pairs of single-inclusive jet and di-jet measurements are available, which focus on different regions of the exchanged boson virtuality Q^2 : a low- Q^2 pair, $5.5 \leq Q^2 \leq 80 \text{ GeV}^2$, and a high- Q^2 pair, $150 \leq Q^2 \leq 15000 \text{ GeV}^2$. The integrated luminosity is, respectively, 290 pb^{-1} and 351 pb^{-1} . On top of the virtuality Q^2 , cross sections are differential in the jet transverse momentum p_T or the di-jet average momentum $\langle p_T \rangle$, respectively for the single-inclusive jet and the di-jet measurements. Massless jets are identified using the k_T algorithm with the R parameter set to $R = 1$. Experimental correlations are available for all measurements, including for points in different single-inclusive jet and di-jet bins at different Q^2 .

The ZEUS measurements correspond to the HERA-I data-taking period, with a centre-of-mass energy of 300 GeV and an integrated luminosity of 38 pb^{-1} , and to the HERA-II data-taking period, with a centre-of-mass energy of 319 GeV and an integrated luminosity of 82 pb^{-1} and 374 pb^{-1} . Cross sections are presented as differential distributions in the vector boson virtuality Q^2 , with $Q^2 \geq 125 \text{ GeV}^2$, and the jet transverse energy E_T or the di-jet average transverse energy $\langle E_T \rangle$, respectively for the single-inclusive jet and the di-jet measurements. Experimental correlations are available across bins within the same set, but not across bins in different datasets. Unlike inclusive DIS structure functions [102], no combination between the H1 and ZEUS results exists to date. A final ZEUS measurement of single-inclusive jet production cross has been published recently [98], however we do not consider it because NNLO QCD corrections to matrix elements are not readily available in a format suitable for fast convolution with PDFs.

Theoretical predictions accurate to NNLO QCD were computed with the NNLOJET code [103, 104] in the zero-mass variable-flavour-number scheme. The central factorisation and renormalisation scales are

$$\mu = \mu_F = \mu_R = Q^2 + (p_T)^2, \quad \mu = \mu_F = \mu_R = Q^2 + \langle p_T \rangle^2, \quad (2.4)$$

respectively for single-inclusive jets and di-jets. Data points for which $\mu \leq 10$ GeV were excluded to ensure that the scale is larger than the b -quark mass. This is necessary because jets are built from massless partons. As in the case of LHC jets, theoretical predictions were released as interpolation grids through the PLOUGHSHARE website [97]. We convert these grids to the PINEAPPL format [69].

3 Methodological framework

In this section, we describe how we quantitatively assess the agreement between the measurements presented in section 2 and the corresponding theoretical predictions for different PDF sets. We first define the figure of merit that we use, and specifically explain how we take into account experimental and theoretical uncertainties in it. We then discuss how this figure of merit may become misleading if the experimental covariance matrix is ill-conditioned, and illustrate how we identify and handle such cases. We finally review the PDF sets that we consider as input for the computation of the theoretical predictions.

3.1 Figure of merit

We quantify the agreement between experimental data and theoretical predictions by computing the (reduced) χ^2 for each dataset

$$\chi^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} (T_i^{(0)} - D_i) (\text{cov}^{-1})_{ij} (T_j^{(0)} - D_j), \quad (3.1)$$

where n_{dat} is the number of data points in the considered dataset, $\{D_i\}$ are the central values of the experimental data, $\{T_i^{(0)}\}$ are the corresponding theoretical predictions, and cov_{ij} is the covariance matrix. Theoretical predictions are evaluated, for both Monte Carlo and Hessian PDF sets, as the prediction obtained with the central PDF $f^{(0)}$, $T_i^{(0)} = T_i(f^{(0)})$. Note that the χ^2 in eq. (3.1) is normalised to the number of data points. Therefore, in case of perfect agreement between data and theory, one expects $\chi^2 \sim 1$, with statistical fluctuations of the order of the standard deviation of the χ^2 distribution, $\sigma_{\chi^2} = \sqrt{2/n_{\text{dat}}}$.

To evaluate eq. (3.1), one needs to compute the covariance matrix cov_{ij} . In addition to experimental uncertainties, one should consider all the relevant sources of theoretical uncertainties, in particular, those associated to missing higher orders (MHO), to PDFs, and to the value of the strong coupling $\alpha_s(m_Z)$. Assuming that all of these theoretical uncertainties follow a Gaussian distribution and that they are mutually independent, they can be incorporated in the covariance matrix following the formalism developed in [105, 106]. Specifically, in this formalism the covariance matrix in eq. (3.1) reads as

$$\text{cov}_{ij} = (\text{cov}_{\text{exp}})_{ij} + (\text{cov}_{\text{th}})_{ij}, \quad (3.2)$$

where the theory covariance matrix is in turn the sum of a MHO, PDF, and α_s contributions

$$(\text{cov}_{\text{th}})_{ij} = (\text{cov}_{\text{mho}})_{ij} + (\text{cov}_{\text{pdf}})_{ij} + (\text{cov}_{\text{as}})_{ij}. \quad (3.3)$$

The experimental covariance matrix is sometimes provided together with the experimental measurements, otherwise, in most cases, we reconstruct it from knowledge of experimental

uncertainties as

$$(\text{cov}_{\text{exp}})_{ij} = \delta_{ij} \sigma_i^{(\text{uncorr})} \sigma_j^{(\text{uncorr})} + \sum_{\ell=1}^{n_{\text{corr}}} \sigma_{i,\ell}^{(\text{corr})} \sigma_{j,\ell}^{(\text{corr})}. \quad (3.4)$$

Here $\sigma_i^{(\text{uncorr})}$ is the sum in quadrature of all the uncorrelated uncertainties, and $\sigma_{i,\ell}^{(\text{corr})}$ is the set of n_{corr} correlated uncertainties. These could be additive or multiplicative, however this distinction is not relevant here, given that eq. (3.1) is only used to quantify the agreement between data and theory. In a fit of PDFs, this distinction is instead relevant because multiplicative uncertainties may bias the determination of the PDF central value and variance. Therefore they would require a specific treatment, by re-defining either the experimental covariance matrix with the t_0 prescription [107] or the figure of merit with additional nuisance parameters [108]. Note that whenever we reconstruct the experimental covariance matrix using eq. (3.4), we implicitly assume that correlated uncertainties are 100% correlated, given that no specific correlation model is provided for the considered datasets, see section 2. Decorrelation remains however possible, using the procedure summarised in section 3.2, and we will actually make use of it, as discussed further below.

The contribution to the covariance matrix due to MHO is evaluated following [105, 106]. Specifically, MHO are estimated as the difference between theoretical predictions computed with fixed and varied renormalisation and factorisation scales, μ_R and μ_F . Several prescriptions defining scale variations are possible. As is common practice in LHC analyses, we adopt the 7-point variation prescription, which gives the MHO covariance matrix

$$(\text{cov}_{\text{mho}})_{ij} = \frac{1}{3} \{ \Delta_i^{+0} \Delta_j^{+0} + \Delta_i^{-0} \Delta_j^{-0} + \Delta_i^{0+} \Delta_j^{0+} + \Delta_i^{0-} \Delta_j^{0-} + \Delta_i^{++} \Delta_j^{++} + \Delta_i^{--} \Delta_j^{--} \}, \quad (3.5)$$

where, for each data point i, j , the shifts are defined as

$$\Delta_i(\kappa_R, \kappa_F) = T_i(\mu_R = \kappa_R \mu_R^{(0)}, \mu_F = \kappa_F \mu_F^{(0)}) - T_i(\mu_R^{(0)}, \mu_F^{(0)}), \quad (3.6)$$

with $(\mu_R^{(0)}, \mu_F^{(0)})$ the central renormalisation and factorisation scales and

$$\begin{aligned} \Delta_i^{+0} &= \Delta_i(2, 1), & \Delta_i^{-0} &= \Delta_i(1/2, 1), & \Delta_i^{0+} &= \Delta_i(1, 1/2), \\ \Delta_i^{0-} &= \Delta_i(1, 1/2), & \Delta_i^{++} &= \Delta_i(2, 2), & \Delta_i^{--} &= \Delta_i(1/2, 1/2). \end{aligned} \quad (3.7)$$

The shifts in the NNLO theory predictions associated to the scale variations, eq. (3.6), are directly evaluated from the PINEAPPL grids [69]. In general, the 7-point MHO theory covariance matrix defined by eq. (3.5) differs from the envelope prescription to estimate MHO uncertainties frequently used in LHC studies.

The contribution to the covariance matrix due to PDF uncertainties is determined, for each of the PDF sets considered (see section 3.3), using the definition of covariance between the random variables T_i and T_j

$$(\text{cov}_{\text{pdf}})_{ij} = \mathbb{E}[(T_i - \mathbb{E}[T_i])(T_j - \mathbb{E}[T_j])], \quad (3.8)$$

where $\mathbb{E}[X]$ denotes the expectation value of the random variable X . For Hessian PDF sets, eq. (3.8) reads

$$(\text{cov}_{\text{pdf}})_{ij} = \sum_{k=1}^{n_{\text{eig}}} (T_i^{(k)} - T_i^{(0)}) (T_j^{(k)} - T_j^{(0)}), \quad (3.9)$$

where $T_i^{(k)} = T_i(f^{(k)})$ is the theoretical prediction computed with the PDF associated to the k -th eigenvalue $f^{(k)}$, and $T_i^{(0)} = T_i(f^{(0)})$ is the theoretical prediction computed with the central PDF $f^{(0)}$. We use the symmetric definition of Hessian PDF uncertainties, since we assume that PDF uncertainties are Gaussian. In case of asymmetric Hessian PDF sets, we replace the difference between $T_{i,j}^{(k)}$ and $T_{i,j}^{(0)}$ in eq. (3.9) with the difference between predictions obtained with positive and negative eigenvectors. For Monte Carlo PDF sets, eq. (3.8) reads

$$(\text{cov}_{\text{pdf}})_{ij} = \frac{1}{n_{\text{rep}}} \sum_{k=1}^{n_{\text{rep}}} \left(T_i^{(k)} - \langle T_i \rangle_{\text{rep}} \right) \left(T_j^{(k)} - \langle T_j \rangle_{\text{rep}} \right), \quad (3.10)$$

where $T_i^{(k)} = T_i(f^{(k)})$ is the theoretical prediction computed with the PDF associated to the k -th replica $f^{(k)}$, and $\langle T_i \rangle_{\text{rep}} = \frac{1}{n_{\text{rep}}} \sum_{k=1}^{n_{\text{rep}}} T_i^{(k)}$ is the average over replicas.

The contribution to the covariance matrix due to the uncertainty of the value of $\alpha_s(m_Z)$ is determined as follows. We take $\alpha_s(m_Z) = 0.118 \pm 0.001$ for all PDF sets considered, consistently with the latest PDG average [87], and we construct

$$(\text{cov}_{\text{as}})_{ij} = \frac{1}{2} \left\{ \Delta_{i,\alpha_s}^+ \Delta_{j,\alpha_s}^+ + \Delta_{i,\alpha_s}^- \Delta_{j,\alpha_s}^- \right\}, \quad (3.11)$$

where, for each data point i, j ,

$$\begin{aligned} \Delta_{i,\alpha_s}^+ &\equiv T_i(\alpha_s = 0.119) - T_i(\alpha_s = 0.118), \\ \Delta_{i,\alpha_s}^- &\equiv T_i(\alpha_s = 0.118) - T_i(\alpha_s = 0.117). \end{aligned} \quad (3.12)$$

The value of α_s in the theory predictions is varied consistently both in the matrix element and in the PDFs, a fact that is streamlined thanks to the usage of PINEAPPL grids. The combination of eq. (3.11) with eq. (3.9) (for a Hessian set) or eq. (3.10) (for a Monte Carlo set) reproduces the prescription of [26], according to which PDF and α_s uncertainties are added in quadrature.

In section 4 we will quantify the agreement between experimental data and theory predictions, obtained with different PDF sets, in terms of the figure of merit given in eq. (3.1). When accounting for all sources of experimental and theoretical uncertainties, we have

$$\chi_{\text{exp+th}}^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} \left(T_i^{(0)} - D_i \right) \left((\text{cov}_{\text{exp}} + \text{cov}_{\text{mho}} + \text{cov}_{\text{pdf}} + \text{cov}_{\text{as}})^{-1} \right)_{ij} \left(T_j^{(0)} - D_j \right), \quad (3.13)$$

with the individual contributions to the covariance matrix combined in quadrature. In order to understand the impact of the various sources of uncertainties entering eq. (3.13), we will also present results for variants of this figure of merit restricted to a subset of the uncertainties, in particular

$$\chi_{\text{exp}}^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} \left(T_i^{(0)} - D_i \right) \left((\text{cov}_{\text{exp}})^{-1} \right)_{ij} \left(T_j^{(0)} - D_j \right), \quad (3.14)$$

which contains only the experimental uncertainties, and

$$\chi_{\text{exp+mho}}^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} \left(T_i^{(0)} - D_i \right) \left((\text{cov}_{\text{exp}} + \text{cov}_{\text{mho}})^{-1} \right)_{ij} \left(T_j^{(0)} - D_j \right), \quad (3.15)$$

defined without the contribution of the PDF and α_s uncertainties. In all cases, the figures of merit are presented normalised to the number of data points of each dataset considered. We emphasise that, when evaluating eq. (3.13), PDFs enter in two different places: through the theory predictions T_i and through the PDF contribution to the total covariance matrix in eq. (3.3).

To further assess the significance of $\chi_{\text{exp+th}}^2$, eq. (3.13), as a measure of the agreement between experimental data and theoretical predictions, we will make use of two additional estimators in section 4. The first estimator is the relative change in the total χ^2 due to the change of input PDF set for a given dataset

$$\Delta\chi^{2(i)} = \frac{\chi_{\text{exp+th}}^{2(i)} - \langle \chi_{\text{exp+th}}^2 \rangle_{\text{pdfs}}}{\langle \chi_{\text{exp+th}}^2 \rangle_{\text{pdfs}}}, \tag{3.16}$$

where the index i runs over the n_{pdfs} input PDF sets considered in the analysis (see section 3.3), and the average over PDF sets is evaluated as

$$\langle \chi_{\text{exp+th}}^2 \rangle_{\text{pdfs}} = \frac{1}{n_{\text{pdfs}}} \sum_{i=1}^{n_{\text{pdfs}}} \chi_{\text{exp+th}}^{2(i)}. \tag{3.17}$$

By construction, $\sum_i \Delta\chi^{2(i)} = 0$. This estimator gauges the relative change in the value of the χ^2 for a given PDF set with respect to the average evaluated over all PDF sets considered. It therefore allows one to disentangle PDF-related effects in the χ^2 from other effects.

The second estimator quantifies the difference of the χ^2 , computed with a given PDF, with respect to the χ^2 averaged over all PDF sets in terms of the number of standard deviations of the χ^2 distribution

$$\Delta n_{\sigma}^{(i)} = \frac{\chi_{\text{exp+th}}^{2(i)} - \langle \chi_{\text{exp+th}}^2 \rangle_{\text{pdfs}}}{\sqrt{2/n_{\text{data}}}}. \tag{3.18}$$

This estimator allows one to compare the χ^2 variation due to the choice of PDF to the expected statistical fluctuations of the χ^2 , and therefore check if this is significant or not. Note indeed that several of the datasets considered contain a relatively small number of data points, so that a large relative change of the χ^2 in eq. (3.16) may be simply explained by large fluctuations due to the small data sample.

3.2 Stability of the experimental covariance matrix

The interpretation of the agreement of theoretical predictions with experimental data, as quantified by the value of the χ^2 , requires some care. As discussed in ref. [109], an inaccurate determination of experimental uncertainty correlations, in otherwise very precise data, may result in an ill-conditioned experimental covariance matrix, which leads in turn to anomalously large values of the χ^2 .

A metric to measure the conditioning of an experimental covariance matrix was introduced in ref. [109], see, in particular, eq. (26). This was defined as the inverse of the smallest singular value of the experimental correlation matrix, and called condition number Z . The value

$(\sqrt{2}Z)^{-1}$ was then demonstrated to be related to the amount by which experimental correlations need to be determined to ensure that the χ^2 remains stable, namely that it does not vary by more than one standard deviation, $\sigma_{\chi^2} = \sqrt{2/n_{\text{dat}}}$. A large value of Z indicates a dataset for which small variations of the correlation model can potentially lead to large χ^2 variations for unchanged data and theory and vice-versa. In ref. [109] a reasonable threshold was defined to be $Z = 4$. This value corresponds to assuming that correlations on uncertainties of the order of a few percent, such as those that affect the LHC measurements considered in this work, be estimated with an absolute uncertainty of less than 0.18. This means that if the correlation between two bins is estimated to be 1.00, while its real value is 0.82, one can expect that the χ^2 deviates from one by more than 1σ even if experimental data and theoretical predictions are perfectly consistent. Note that the smaller the uncertainty, the higher the required precision with which correlations need to be known to be within a 1σ variation of the χ^2 . As explicitly shown in ref. [109] as part of a toy model (see in particular eq. (29) and figure 3), for a 1% (2.5%) uncertainty, correlations ought to be known with a precision of 2% (12%). The choice $Z = 4$ (the same for all experiments) should therefore be seen as a practical diagnostic choice.

In some cases, a large value of Z may not imply a pathological behaviour of the experimental data. A typical case is the one in which the luminosity uncertainty, which by definition is 100% correlated across all bins of a given dataset, is the largest of all uncertainties. In this case, we expect the condition number Z to be large. The same would happen with any experimental uncertainty that is fully correlated across bins for specific experimental reasons. In these cases, we should compute Z upon excluding these uncertainties from the reconstruction of the experimental covariance matrix. For the sake of this work, we single out only the luminosity uncertainty as 100% correlated, given that we do not have complete knowledge of which other uncertainties are also undoubtedly 100% correlated. We then split the experimental covariance matrix into two components, one that contains only the luminosity uncertainty, and one that contains all of the other uncertainties. We compute Z , which we call $Z_{\mathcal{L}}$, for the latter covariance matrix and regularise it, if necessary; we then compute the χ^2 using the sum of the original luminosity covariance matrix and the regularised covariance matrix. Clearly, this procedure might decorrelate systematic uncertainties that owe to be 100% correlated, or decorrelate too much some other systematic uncertainties. Nevertheless, we consider that the procedure remains a useful diagnosis tool when information on specific decorrelation models, provided under the experimental guidance, are not available. In this respect, we shall also note that here our aim is not to characterise a dataset for inclusion (or not) in a PDF determination, but rather to comparatively assess the ability of various PDF sets to describe the data. We consider that the application of our regularisation procedure does not alter our judgement on such an ability (see appendix A).

An alternative estimator to assess the conditioning of the experimental correlation matrix, sometimes used in experimental analyses, is λ_ρ , defined as the ratio of the smallest to the largest eigenvalues of the experimental correlation matrix. A small value of λ_ρ indicates a large spread of eigenvalues, with the directions associated to the smallest ones almost degenerate. These degeneracies are those that lead to a ill-conditioned matrix.

In table 2 we display, for each dataset listed in table 1 and separately for each observable, the number of data points, n_{dat} , and the condition numbers λ_ρ , Z , and $Z_{\mathcal{L}}$. For normalised

distributions $Z = Z_{\mathcal{L}}$ by construction. For datasets which do not provide the breakdown of systematic uncertainties but instead only the overall covariance matrix, $Z_{\mathcal{L}}$ is computed by subtracting from this covariance matrix a covariance matrix constructed only from the 100%-correlated luminosity uncertainty. In the case of the CMS top-quark pair distribution, this procedure is however not applied, given that the measurement is the combination of events recorded with different luminosities. We therefore leave the corresponding entry blank in table 2. Whenever a dataset is presented in different variants, for example as absolute or normalised distributions or for two different values of the jet radius R , we indicate with a (*) the one used in section 4. We select the distributions that feature the lowest value of $Z_{\mathcal{L}}$.

The values of the condition numbers λ_{ρ} and Z reported in table 2 consistently indicate that the experimental correlation and covariance matrices are ill-conditioned for a subset of the analyzed datasets, according to the criterion of refs. [8, 109] ($Z > 4$). For some of them, such as the ATLAS $d\sigma^Z/d|y_{\ell\bar{\ell}}|$ measurement at 8 TeV, and to a lesser extent for LHCb $d\sigma^Z/dy_{\ell\bar{\ell}}$, this high Z value is explained by the dominance of the luminosity uncertainty: in these cases, $Z_{\mathcal{L}}$ is indeed significantly smaller than Z . For all the other datasets, $Z \sim Z_{\mathcal{L}}$. Relatively high values of Z are found for the ATLAS and CMS single-inclusive jet and di-jet datasets, a fact that was already observed in the case of the corresponding measurements at 8 TeV, for which various decorrelation models have been proposed and tested [10, 91, 109–111]. We finally observe that the value of Z can fluctuate by a large amount across different differential measurements in the same dataset. For instance, the 13 TeV ATLAS $t\bar{t}$ hadronic dataset provides single-differential distributions in $m_{t\bar{t}}$ and in $|y_{t\bar{t}}|$, associated to values of Z respectively of 64.5 and 1.77.

The χ^2 of the datasets listed in table 2 will therefore need to be interpreted with care, in particular taking into account the possibility that it be spuriously high due to a misestimate of experimental correlations. To avoid this issue, in section 4 we will compute the χ^2 upon regularisation of the experimental covariance matrix, for all the datasets with $Z_{\mathcal{L}} > 4$. We use the procedure laid out in ref. [109]. This procedure consists in clipping the singular values of the correlation matrix to a constant, whenever these are smaller than that, while leaving the rest of the singular vectors unchanged. This way, directions that do not contribute to instability are not affected and the alteration to the original matrix is minimal. The clipping constant is chosen to be $\delta^{-1} = Z$, where the value of $Z = 4$ was determined empirically in ref. [109]. The values of the χ^2 computed with the unregularised experimental covariance matrix are collected in appendix A.

3.3 PDF sets

The computation of the theoretical predictions that enter the χ^2 require a choice of PDFs as input. In this work, we consider the following PDF sets: ABMP16 [5], CT18, CT18A, and CT18Z [6], MSHT20 [7], NNPDF3.1 [112], NNPDF4.0 [8], PDF4LHC15 [19], and PDF4LHC21 [26]. These PDF sets are the most widely used by LHC experimental collaborations in their analyses. The main features of each of them are summarised as follows.

ABMP16 [5]. This PDF determination is based on DIS, Drell-Yan, single top and top-quark pair production measurements. The underlying theory calculations are based on a Fixed Flavour Number (FFN) scheme, with $n_f = 3, 4, 5$. The strong coupling constant

Process	n_{dat}	λ_ρ	Z	$Z_{\mathcal{L}}$
ATLAS 13 TeV Z $1/\sigma d\sigma/dp_T^{\ell\bar{\ell}}$	38	1.9×10^{-1}	1.10	1.10
CMS 13 TeV W^+ $d\sigma/d \eta $	18	8.3×10^{-5}	25.1	19.0
CMS 13 TeV W^- $d\sigma/d \eta $	18	8.9×10^{-5}	26.0	18.0
LHCb 13 TeV Z $d\sigma/dy^Z$	17	1.9×10^{-3}	5.92	2.09
ATLAS 8 TeV Z $d\sigma/d y $	7	3.2×10^{-4}	21.6	2.10
ATLAS 13 TeV $t\bar{t}$ all hadr. $d\sigma/dm_{t\bar{t}}$ (*)	9	2.4×10^{-3}	7.27	7.24
ATLAS 13 TeV $t\bar{t}$ all hadr. $1/\sigma d\sigma/dm_{t\bar{t}}$	9	3.9×10^{-5}	64.7	64.7
ATLAS 13 TeV $t\bar{t}$ all hadr. $d\sigma/d y_{t\bar{t}} $	12	3.3×10^{-3}	5.27	5.25
ATLAS 13 TeV $t\bar{t}$ all hadr. $1/\sigma d\sigma/d y_{t\bar{t}} $ (*)	12	8.9×10^{-2}	1.77	1.77
ATLAS 13 TeV $t\bar{t}$ all hadr. $d^2\sigma/dm_{t\bar{t}} d y_{t\bar{t}} $ (*)	11	4.4×10^{-3}	4.83	4.81
ATLAS 13 TeV $t\bar{t}$ all hadr. $1/\sigma d^2\sigma/dm_{t\bar{t}} d y_{t\bar{t}} $	11	9.4×10^{-5}	52.1	52.1
ATLAS $t\bar{t} \ell + j$ $d\sigma/dm_{t\bar{t}}$	9	5.2×10^{-4}	16.2	15.9
ATLAS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/dm_{t\bar{t}}$ (*)	9	3.0×10^{-3}	7.62	7.62
ATLAS 13 TeV $t\bar{t} \ell + j$ $d\sigma/dp_T^t$	8	5.8×10^{-4}	16.8	16.6
ATLAS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/dp_T^t$ (*)	8	2.5×10^{-3}	8.46	8.46
ATLAS 13 TeV $t\bar{t} \ell + j$ $d\sigma/d y_t $	5	1.5×10^{-3}	11.7	11.5
ATLAS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/d y_t $ (*)	5	9.6×10^{-2}	2.06	2.06
ATLAS 13 TeV $t\bar{t} \ell + j$ $d\sigma/d y_{t\bar{t}} $	7	6.2×10^{-4}	15.7	15.4
ATLAS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/d y_{t\bar{t}} $ (*)	7	7.8×10^{-2}	2.26	2.26
CMS 13 TeV $t\bar{t} \ell + j$ $d\sigma/dm_{t\bar{t}}$	15	1.1×10^{-2}	3.90	—
CMS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/dm_{t\bar{t}}$ (*)	15	3.0×10^{-2}	3.51	3.51
CMS 13 TeV $t\bar{t} \ell + j$ $d\sigma/dp_T^t$	16	7.5×10^{-3}	4.04	—
CMS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/dp_T^t$ (*)	16	1.3×10^{-1}	1.78	1.78
CMS $t\bar{t} \ell + j$ $d\sigma/d y_t $	11	3.3×10^{-3}	5.75	—
CMS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/d y_t $ (*)	11	2.7×10^{-1}	1.36	1.36
CMS 13 TeV $t\bar{t} \ell + j$ $d\sigma/d y_{t\bar{t}} $	10	1.2×10^{-3}	9.68	—
CMS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d\sigma/d y_{t\bar{t}} $ (*)	10	1.9×10^{-1}	1.53	1.53
CMS 13 TeV $t\bar{t} \ell + j$ $d^2\sigma/dm_{t\bar{t}} d y_{t\bar{t}} $	35	8.1×10^{-5}	22.4	—
CMS 13 TeV $t\bar{t} \ell + j$ $1/\sigma d^2\sigma/dm_{t\bar{t}} d y_{t\bar{t}} $ (*)	35	1.8×10^{-4}	17.2	17.2
ATLAS 13 TeV single-inclusive jets $d^2\sigma/dp_T d y $	177	2.6×10^{-5}	16.9	16.2
CMS 13 TeV single-inclusive jets ($R = 0.4$) $d^2\sigma/dp_T d y $ (*)	78	1.1×10^{-4}	13.3	13.1
CMS 13 TeV single-inclusive jets ($R = 0.7$) $d^2\sigma/dp_T d y $	78	9.0×10^{-5}	14.8	14.5
ATLAS 13 TeV di-jets $d^2\sigma/dm_{jj} d y^* $	136	3.8×10^{-5}	16.8	15.6
H1 single-inclusive-jets (low Q^2) $d^2\sigma/dQ^2 dp_T$	48	7.6×10^{-3}	6.00	5.91
H1 single-inclusive-jets (high Q^2) $d^2\sigma/dQ^2 dp_T$	24	7.0×10^{-3}	1.46	1.19
ZEUS single-inclusive jets (low luminosity) $d^2\sigma/dQ^2 dE_T$	30	5.0×10^{-2}	1.87	1.82
ZEUS single-inclusive jets (high luminosity) $d^2\sigma/dQ^2 dE_T$	30	1.9×10^{-2}	2.56	2.43
H1 di-jets (low Q^2) $d^2\sigma/dQ^2 d\langle p_T \rangle$	48	9.0×10^{-2}	7.67	7.42
H1 di-jets (high Q^2) $d^2\sigma/dQ^2 d\langle p_T \rangle$	24	1.0×10^{-1}	1.60	1.45
ZEUS di-jets $d^2\sigma/dQ^2 d\langle E_T \rangle$	22	1.5×10^{-2}	2.83	2.72

Table 2. The number of data points, n_{dat} , the condition numbers λ_ρ , Z , and $Z_{\mathcal{L}}$ for all datasets considered, see the text for their definition. When the $Z_{\mathcal{L}}$ estimator cannot be unambiguously computed (as explained in the text) the corresponding entry is left blank. Whenever different variants or distributions exist for a dataset, we indicate with a (*) the one used in section 4.

is determined alongside the PDFs yielding $\alpha_s(m_Z) = 0.1147 \pm 0.0008$ with $n_f = 5$, though a variant with a fixed value $\alpha_s(m_Z) = 0.118$ is also provided. The PDFs are parametrised at the input scale $Q_0 = 1 \text{ GeV}$ with a fixed functional form. The charm PDF is assumed to be purely perturbative, therefore it is generated by partonic DGLAP evolution above the charm quark mass, whose value is a parameter of the fit. Hessian symmetric PDF uncertainties are determined from variations $\Delta\chi^2 = 1$.

CT18 [6]. The CT18 family of PDF determinations is based on DIS, Drell-Yan, single-inclusive jet, and top-quark pair production measurements. The underlying theory calculations are based on a General Mass Variable Flavour Number (GM-VFN) scheme, specifically ACOT- χ [113–116], and use a fixed value of the strong coupling as input. Parton distributions are parametrised at the input scale $Q_0 = 1.3 \text{ GeV}$, equal to the charm pole mass m_c^{pole} , in terms of Bernstein polynomials, the charm PDF is purely perturbative, and Hessian symmetric PDF uncertainties are determined by means of a dynamical tolerance factor $\Delta\chi^2 > 1$. The ATLAS 7 TeV W/Z data [117] is not included in the default CT18 PDF set. Alternate sets are determined including this dataset (CT18A), a new scale choice for low- x DIS data (CT18X), or all of the above with a slightly higher value of the charm mass (CT18Z).

MSHT20 [7]. This PDF determination is based on DIS, Drell-Yan, Drell-Yan with jet, single-inclusive jet, and top-quark pair production measurements. The fit is based on the Thorne-Roberts variant of the GM-VFN scheme [118], and uses a fixed value of the strong coupling as input. Parton distributions are parametrised at the input scale $Q_0 = 1 \text{ GeV}$ in terms of Chebyshev polynomials, the charm PDF is purely perturbative (with charm pole mass $m_c^{\text{pole}} = 1.4 \text{ GeV}$), and Hessian symmetric uncertainties are determined by means of a dynamical tolerance factor $\Delta\chi^2 > 1$.

NNPDF3.1 [28]. This PDF determination is based on DIS, Drell-Yan, Drell-Yan with jet, single-inclusive jet, and top-quark pair production measurements. The fit is based on the FONLL GM-VFN scheme [119] and uses a fixed value of the strong coupling constant as input. Parton distributions are parametrised at the initial scale $Q_0 = 1.65 \text{ GeV}$ in terms of deep neural networks, optimised by means of a genetic algorithm. The charm PDF is fitted on the same footing as lighter quark flavours (with a charm pole mass $m_c^{\text{pole}} = 1.51 \text{ GeV}$). PDF uncertainties are determined from a Monte Carlo sampling of experimental uncertainties.

NNPDF4.0 [8]. This PDF determination is based on DIS, Drell-Yan, Drell-Yan with jet, single-inclusive jet and di-jet, single top and top-quark pair, and prompt photon production measurements. The fit is based on the same treatment of quark masses, running coupling, charm quark PDF, and uncertainty representation as NNPDF3.1. In comparison to NNPDF3.1, NNPDF4.0 is however characterised by several methodological differences: newer theoretical constraints, in particular on PDF positivity and integrability, are implemented; PDFs are parametrised with a single neural network, optimised by means of gradient descent; hyperparameters, such as those that define the architecture

of the neural network, are determined by means of an automated scan of the space of models that selects the optimal one [8, 120]; and the methodology is closure tested [121].

PDF4LHC15 [19]. This PDF set is the Monte Carlo combination of the CT14 [122], MMHT2014 [123], and NNPDF3.0 [124] PDF sets. The combination is performed by first converting the CT14 and MMHT2014 Hessian PDF sets into Monte Carlo PDF sets by means of the algorithm developed in [20, 21]. For each of the three PDF sets 300 Monte Carlo replicas are generated, that are subsequently collated in a single set. The number of replicas is finally reduced by means of the compression algorithm developed in [25] or converted to a single Hessian set by means of the algorithm developed in [23].

PDF4LHC21 [26]. This PDF set is the Monte Carlo combination of the CT18', MSHT20, and NNPDF3.1' PDF sets. The CT18' and NNPDF3.1' PDF sets are variants of the CT18 and NNPDF3.1 PDF sets: both of these differ from the corresponding baseline sets for the values of the charm and bottom quark pole masses, which are set to values common to those used in MSHT20, $m_c^{\text{pole}} = 1.4 \text{ GeV}$ and $m_b^{\text{pole}} = 4.75 \text{ GeV}$. The NNPDF3.1' PDF set differs from NNPDF3.1 for a number of additional variations in the input dataset and in the details of the theoretical computations, see section 2.3 in [26]. The combination is carried out as in PDF4LHC15.

In all cases, we use PDF sets accurate to NNLO with a common, fixed value of $\alpha_s(m_Z) = 0.118$. Note that NNLO corrections to hadronic processes were included in all of the aforementioned PDF sets by means of K -factors, whereas here we make predictions by means of exact NNLO computations. This fact is however immaterial, given the very weak dependence of K -factors on PDFs [125]. In the case of ABMP16, we use the set with $n_f = 5$ active flavours. For ABMP16, CT18, and MSHT20, we consider Hessian sets; for NNPDF3.1, NNPDF4.0, PDF4LHC15, and PDF4LHC21, we consider Monte Carlo sets composed of 100 replicas. In figure 1 we compare the partonic luminosities, defined by eqs. (1–4) of [126], obtained with the ABMP16, CT18, MSHT20, NNPDF4.0, and PDF4LHC21 PDF sets. Results are displayed as a function of the invariant mass of the final state m_X at a centre-of-mass energy $\sqrt{s} = 13 \text{ TeV}$ and are normalised to PDF4LHC21. Comparison using other PDF sets can be seen in [3].

We do not consider PDF sets including QED corrections [127–129], aN³LO QCD corrections [43, 130] or MHOU [106], the reason being that these are not commonly used in LHC experimental analyses. This said, the computation of the χ^2 does not change if one uses any of these PDF sets. We will study the phenomenological implications of QED, aN³LO, and of MHOU corrections to the PDFs in the appraisal of LHC data in future work. An exception is represented by the high-precision ATLAS 8 TeV inclusive dilepton rapidity measurement [42], for which predictions based on the NNPDF4.0 QED [129], MHOU [106], and aN³LO [130] PDF sets will be considered in section 4.2.

4 Data-theory comparison appraisal

In this section, we quantify the agreement between the experimental data and the corresponding theoretical predictions presented in section 2 according to the estimators and upon

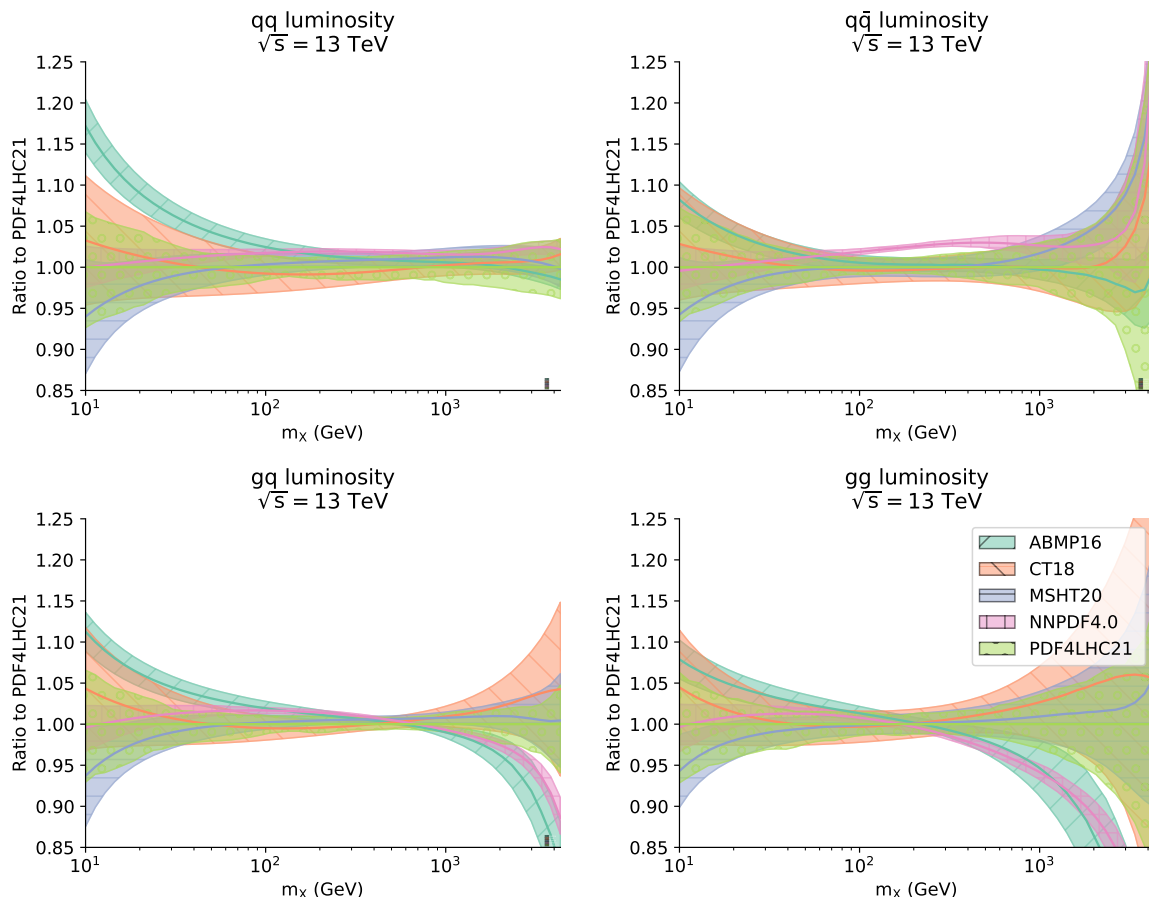


Figure 1. The quark-quark (top left), quark-antiquark (top right), gluon-quark (bottom left), and gluon-gluon (bottom right) partonic luminosities, eqs. (1–4) of [126], as a function of the invariant mass of the final state m_X at a centre-of-mass energy $\sqrt{s} = 13$ TeV obtained with the ABMP16, CT18, MSHT20, NNPDF4.0, and PDF4LHC21 PDF sets. Results are normalised to PDF4LHC21.

variations of the input PDF sets discussed in section 3. We examine datasets for each process in turn. For each of these, we provide three complementary ways of visualizing the data-theory agreement: a table with the values of $\chi^2_{\text{exp+th}}$, and χ^2_{exp} , eqs. (3.13) and (3.14), evaluated with all the PDF sets summarised in section 3.3;¹ a set of histograms in which the total $\chi^2_{\text{exp+th}}$, eq. (3.13), is split into the components $\chi^2_{\text{exp+mho}}$, eq. (3.15), and $\chi^2_{\text{exp+th}}$, eq. (3.14), albeit only for CT18, MSHT20, NNPDF4.0, and PDF4LHC21; and a set of data-theory comparison plots, only for NNPDF4.0 and PDF4LHC21, in which the PDF+ α_s and MHO uncertainties are displayed separately for selected data points. For all measurements with $Z_{\mathcal{L}} > 4$ (see table 2), the experimental covariance matrix is regularised as explained in section 3.2. We finally provide a collective visualisation of the $\Delta\chi^{2(i)}$ and $\Delta n_{\sigma}^{(i)}$ estimators, eqs. (3.17) and (3.18), for the CT18, MSHT20, NNPDF4.0, and PDF4LHC21 PDF sets. The values of $\chi^2_{\text{exp+th}}$ obtained without regularisation of the experimental covariance matrix are given, for the subset of measurements with $Z_{\mathcal{L}} > 4$, in appendix A. Additional histogram

¹For ease of reference, in this table we also provide the number of data points n_{dat} and the standard deviation of the χ^2 per point distribution $\sqrt{2/n_{\text{dat}}}$.

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
ATLAS 13 TeV $Z \frac{1}{\sigma} \frac{d\sigma}{dp_T^{\ell}}$	38	0.23	$\chi_{\text{exp+th}}^2$	0.36	0.31	0.42	0.59	0.40	0.39	0.50	0.31	0.38
			χ_{exp}^2	0.80	1.18	2.38	4.91	1.58	1.20	2.20	0.83	1.64
CMS 13 TeV $W^+ \frac{d\sigma}{d \eta }$	18	0.33	$\chi_{\text{exp+th}}^2$	1.31	1.20	1.11	1.06	1.26	0.85	0.96	1.15	0.98
			χ_{exp}^2	1.41	1.67	1.30	1.31	1.37	0.97	1.12	1.38	1.27
CMS 13 TeV $W^- \frac{d\sigma}{d \eta }$	18	0.33	$\chi_{\text{exp+th}}^2$	1.56	1.15	1.11	1.10	1.43	1.12	1.60	1.14	1.20
			χ_{exp}^2	1.60	1.89	1.43	1.38	1.57	1.64	1.95	1.54	1.54
LHCb 13 TeV $Z \frac{d\sigma}{dyZ}$	18	0.33	$\chi_{\text{exp+th}}^2$	2.14	2.19	2.26	2.08	2.28	2.21	2.26	2.15	2.07
			χ_{exp}^2	2.28	3.09	2.91	2.62	2.66	2.70	2.48	3.06	2.67

Table 3. The values of $\chi_{\text{exp+th}}^2$, eq. (3.13), and of χ_{exp}^2 , eq. (3.14), for the ATLAS, CMS, and LHCb Drell-Yan gauge boson production measurements at the LHC 13 TeV of table 1, computed with each of the PDF sets summarised in section 3.3. The experimental covariance matrix of the CMS dataset is regularised as explained in section 3.2. The unregularised values of χ_{exp}^2 are collected in table 9 of appendix A.

and data-theory comparison plots, for the subset of measurements not highlighted in this section, are given in appendix B.

4.1 Drell-Yan weak boson production measurements at 13 TeV

We start by considering the three LHC Drell-Yan weak boson production measurements at a centre-of-mass energy of 13 TeV outlined in section 2. The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$, computed with each of the PDF sets summarised in section 3.3, are reported in table 3. The experimental covariance matrix of the CMS dataset is regularised as explained in section 3.2, see appendix A for the unregularised values of $\chi_{\text{exp+th}}^2$. The breakdown of $\chi_{\text{exp+th}}^2$ into $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 is displayed in figure 2. The data-theory comparison is displayed in figure 3. Each plot consists of three panels: the upper one displays the measured and predicted cross sections, with experimental and total (MHO and PDF+ α_s) theoretical uncertainties; the middle one displays the same cross sections normalised to the experimental central value; the lower one displays the relative PDF+ α_s and MHO uncertainties separately. Experimental error bars correspond to the total uncorrelated uncertainty. Correlated uncertainties are included by shifting the central experimental value, by an amount determined as explained in appendix B of [108].

From inspection of table 3 and of figure 2, we observe that the values of $\chi_{\text{exp+th}}^2$, computed with different input PDFs, are generally closer to each other than the corresponding values of χ_{exp}^2 . This fact suggests that the inclusion of theory uncertainties is essential to assess the predictive power of a given PDF set. Moreover, the values of $\chi_{\text{exp+th}}^2$ are very similar across PDF sets: this is manifest in the case of the ATLAS and LHCb datasets, and true on average for the CMS dataset. In this latter case, the PDF sets with larger values of $\chi_{\text{exp+th}}^2$ on the W^+ dataset have the smaller values of $\chi_{\text{exp+th}}^2$ on the W^- dataset, and the other way around. For the CMS W^+ distribution, whereas the total $\chi_{\text{exp+th}}^2$ remains within 1σ of the

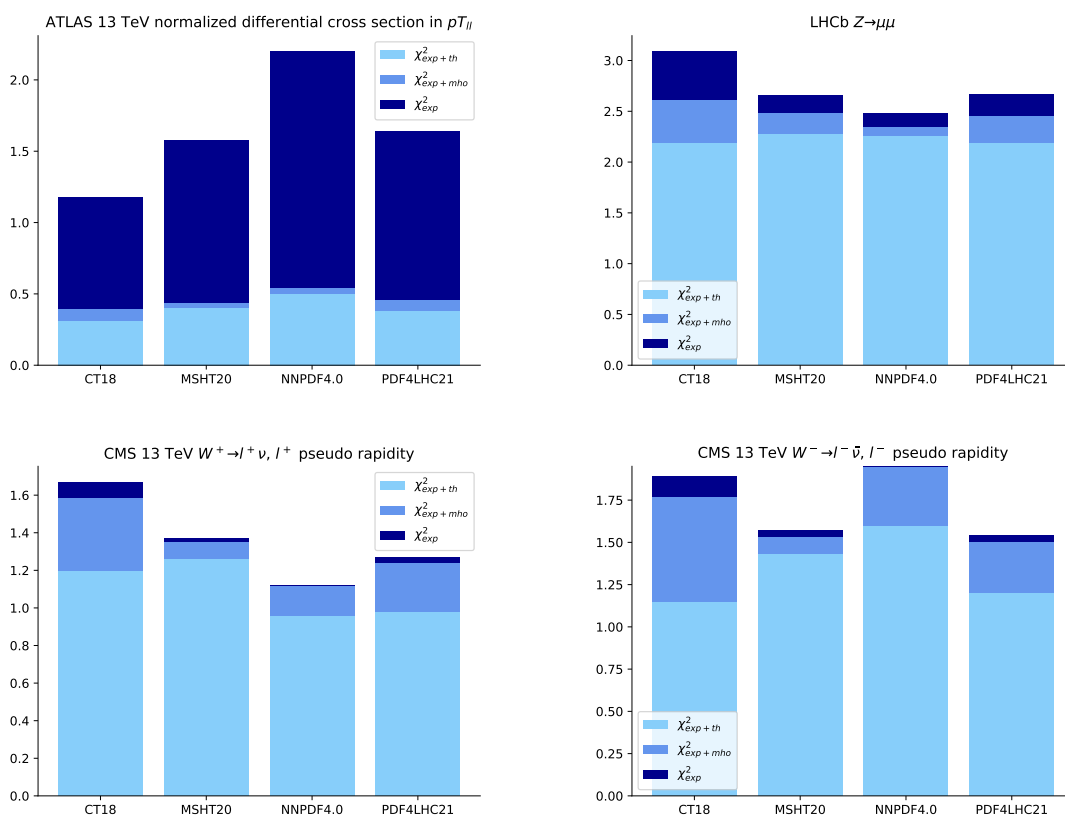


Figure 2. The breakdown of χ^2_{exp+th} , eq. (3.13), into $\chi^2_{exp+mho}$, eq. (3.15), and χ^2_{exp} , eq. (3.14), for the ATLAS ($n_{dat} = 38$, $\sqrt{2}/n_{dat} = 0.23$), CMS (each set made of $n_{dat} = 18$, $\sqrt{2}/n_{dat} = 0.23$), and LHCb ($n_{dat} = 18$, $\sqrt{2}/n_{dat} = 0.33$) Drell-Yan gauge boson production measurements at the LHC 13 TeV. Note that the inclusion of MHOUs has a negligible impact on the χ^2 values of the CMS 13 TeV W distributions when theoretical predictions are computed with NNPDF4.0, hence the dark blue component of the corresponding histogram is hardly visible.

χ^2 distribution per unit point for all PDF sets, the experimental χ^2_{exp} is close to one only for NNPDF3.1 and NNPDF4.0. In the case of the CMS W^- distribution, the ABMP16 and NNPDF4.0, and to a lesser extent the MSHT20 set do yield a somewhat worse description, in that the corresponding values of χ^2_{exp+th} are almost 2σ away from the unit expectation. For the purely experimental χ^2_{exp} , the predictions obtained with NNPDF4.0 and CT18 are almost 3σ away from one. Despite these differences, when all uncertainties are kept into account, we cannot single out a PDF set that, overall, generalises better than another on these datasets.

The breakdown of χ^2_{exp+th} into its theoretical components depends on the dataset and on the PDF set. The component due to MHO, gauged from the difference between χ^2_{exp} and $\chi^2_{exp+mho}$, dominates the ATLAS measurement, irrespective of the PDF set, whereas it is less prominent in the other datasets. For CMS, this is almost immaterial, irrespective of the PDF set. For LHCb, irrespective of the PDF set, this is typically as large as the component due to PDF+ α_s uncertainties, gauged from the difference between $\chi^2_{exp+mho}$ and χ^2_{exp+th} . This latter component may depend on the PDF set, being usually larger for PDF sets affected

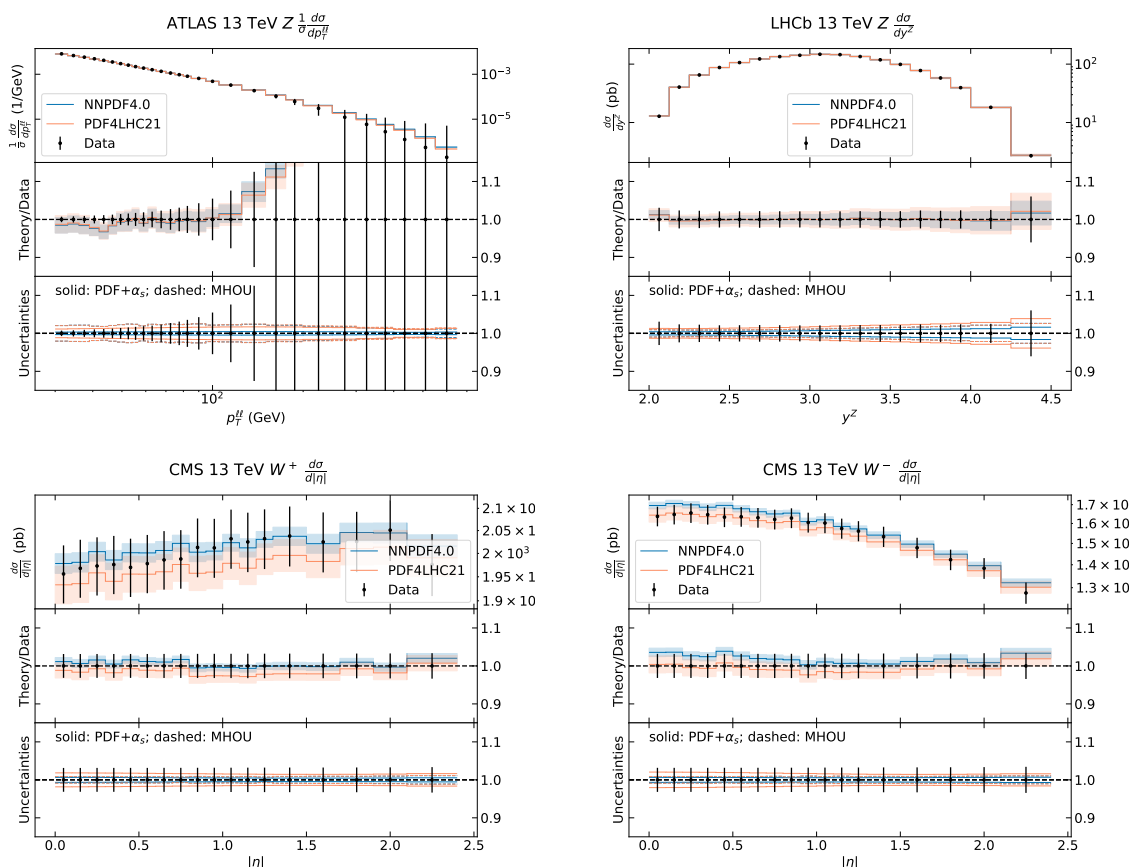


Figure 3. Data-theory comparison for the ATLAS, CMS, and LHCb Drell-Yan gauge boson production measurements at the LHC 13 TeV of table 1. (Upper panels) The measured and predicted cross sections, with experimental and total (MHO and PDF+ α_s) theoretical uncertainties. (Middle panels) The same cross sections normalised to the experimental central value. (Lower panels) The relative PDF+ α_s (dashed) and MHO (solid) uncertainties separately. In all panels, the experimental error bars correspond to the total uncorrelated uncertainty. Correlated uncertainties are kept into account by shifting the central experimental value as explained in appendix B of [108].

by the largest uncertainties, such as CT18 and PDF4LHC21, see figure 1. All these facts are a consequence of how the various partonic channels contribute to the cross sections of these processes. The ATLAS measurement receives its leading contribution, which is $\mathcal{O}(\alpha_s)$, from the quark-gluon partonic luminosity. The CMS and LHCb measurements receive their leading contributions, which are $\mathcal{O}(\alpha_s^0)$, from quark-antiquark partonic luminosities, yet in different regions of x , given that they are central and forward rapidity measurements: the former at intermediate values of x ; the latter at large values of x .

The quality of the data description is generally good, being $\chi_{\text{exp+th}}^2 \sim 1$, except for LHCb, for which $\chi_{\text{exp+th}}^2 \sim 2$, irrespective of the PDF set. For ATLAS, $\chi_{\text{exp+th}}^2 \sim 0.4$, again irrespective of the PDF set. This value is anomalously small, and may point towards the fact that MHO are actually overestimated by scale variations. Discrepancies between data and theory that may lead to these results are seen in figure 3, where the alignment of experimental data and theoretical predictions is optimal, within their uncertainties, for

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
ATLAS 8 TeV Z $\frac{d\sigma}{d y }$	7	0.53	$\chi^2_{\text{exp+th}}$	4.25	1.52	1.52	1.18	1.37	1.61	3.83	1.23	1.09
			χ^2_{exp}	7.36	14.0	4.63	4.31	2.14	4.70	7.90	7.41	1.93

Table 4. Same as table 3 for the ATLAS Drell-Yan gauge boson production measurements at the LHC 8 TeV [42].

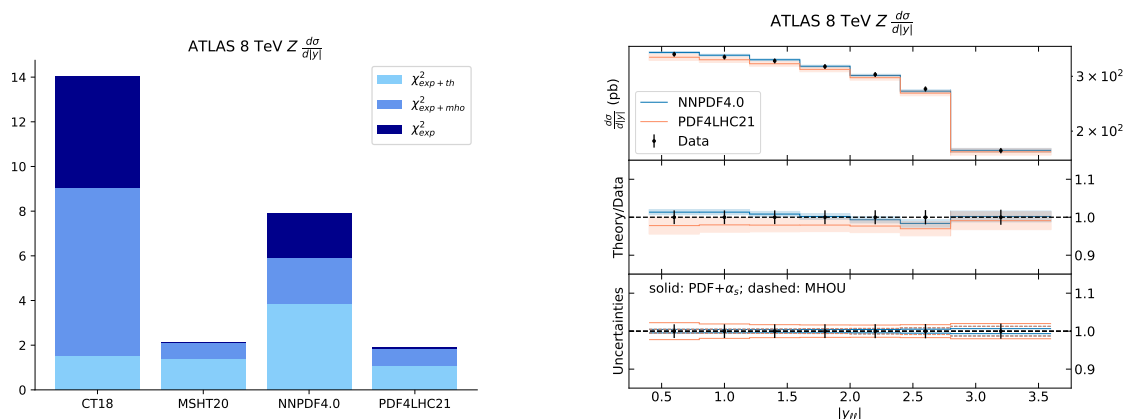


Figure 4. Same as figures 2 (left) and 3 (right) for the ATLAS Drell-Yan gauge boson production measurements at the LHC 8 TeV [42] ($n_{\text{dat}} = 7$, $\sqrt{2/n_{\text{dat}}} = 0.53$).

all datasets. We therefore conclude that the somewhat high $\chi^2_{\text{exp+th}}$ for LHCb is due to experimental correlations, and will likely decrease once the dataset is included in a fit. Note finally that the quality of the data description of the CMS measurement would have been rather worse, at face value, had the regularisation procedure described in section 3.2 not been applied. The values of $\chi^2_{\text{exp+th}}$ obtained without it are reported in appendix A. As we can see from figure 3, theoretical predictions are almost spot on experimental measurements. The otherwise very large values of the χ^2 obtained without regularisation are spurious, and denote an ill-conditioning of their experimental covariance matrix.

4.2 The ATLAS 8 TeV inclusive Z boson production measurement

We then consider the ATLAS measurement of Drell-Yan Z boson production at the LHC 8 TeV outlined in section 2. The values of χ^2_{exp} and $\chi^2_{\text{exp+th}}$, computed with each of the PDF sets summarised in section 3.3, are reported in table 4. The breakdown of $\chi^2_{\text{exp+th}}$ into $\chi^2_{\text{exp+mho}}$ and χ^2_{exp} and the data-theory comparison are displayed in figure 4, in the same format as figures 2 and 3.

From inspection of table 4 and figure 4, we observe that the values of $\chi^2_{\text{exp+th}}$ decrease significantly with respect to χ^2_{exp} . As already remarked for the other Drell-Yan data, this fact further indicates that a careful account of theoretical uncertainties is crucial to assess the predictive power of a PDF set. For CT18 and NNPDF4.0, the MHO and PDF+ α_s

contributions to the χ^2 have approximately the same size, and are relatively large. For MSHT20 and PDF4LHC21, the MHO contribution to the χ^2 is essentially immaterial. This is possibly due to the fact that there is a large variance in the quality of the description of this dataset before including theoretical uncertainties in the computation of the χ^2 : even if all PDF sets provide an unsatisfactory description of the data, MSHT20 and PDF4LHC21 have a χ_{exp}^2 of order 2, whereas all of the others have a χ_{exp}^2 of order 5–10. Once theoretical uncertainties are included, one gets $\chi_{\text{exp+th}}^2$ of the order of 1, except for ABMP16 and NNPDF4.0, for which $\chi_{\text{exp+th}}^2$ is equal to 3.47 and 3.83. The discrepancy between experimental data and theoretical predictions obtained with NNPDF4.0 instead of PDF4LHC21 is visible in the right panel of figure 4. The shape of the NNPDF4.0 prediction displays a peculiar dip around a value of the dilepton rapidity of 2.7.

This is a case in which NNPDF4.0 seems to perform poorly. This fact is a little surprising because, as discussed in section 2.1, an earlier version [68] of the ATLAS measurement of [42] was included in NNPDF4.0. This is a substantial difference with respect to all of the other measurements examined in this work, which, albeit corresponding to production processes already included in NNPDF4.0, are completely out of sample. For this reason, we consider that, only in this case, some more extensive investigations are needed. To this purpose, we recompute the values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$ using the NNPDF4.0 PDF sets that include QED corrections [129], MHOUs [106], and aN³LO corrections and MHOUs [130]. All these PDF sets include the ATLAS Drell-Yan Z boson production measurements at 8 TeV presented in [67, 68]. Furthermore, to understand the interplay of these measurements with the new version considered here [42] (see section 2 for details), we perform the following additional fits:

- (a) a fit equivalent to the NNLO NNPDF4.0 baseline fit excluding the ATLAS measurement of [68];
- (b) a fit equivalent to the NNLO NNPDF4.0 baseline fit in which the ATLAS measurement of [68] is replaced with that of [42];
- (c) a fit equivalent to fit (b), in which the ATLAS measurement of [42] is weighted as explained in section 4.2.3 of [8];
- (d) a fit equivalent to fit (b), including MHOUs.

The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$ computed with the baseline NNPDF4.0 PDF set and with all the aforementioned PDF sets are collected in table 5. Values are displayed for the ATLAS measurement of [42], which is included only in fits (b), (c), and (d), for the ATLAS measurement of [68], which is included in the NNPDF4.0, aN³LO MHOUs, MHOUs, and QED fits, and for the subset of the ATLAS measurement of [68] corresponding to the invariant mass bin of the Z peak. This way, the kinematic coverage is the same as in [42]. The results corresponding to the NNPDF4.0 baseline fit are the same as in table 4.

From table 5, we make the following conclusions. The ATLAS dataset of [68] is described fairly well by NNPDF4.0, including the bin corresponding to the Z -peak invariant mass, whereas the dataset of [42] is not, even when accounting for theoretical uncertainties in the computation of $\chi_{\text{exp+th}}^2$. This state of affairs does not change if one considers variants of the NNPDF4.0 PDF sets including N³LO corrections, MHOUs, or QED corrections. It is

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		NNPDF4.0	aN^3LO	MHOU	QED	fit (a)	fit (b)	fit (c)	fit (d)
ATLAS 8 TeV $Z \frac{d\sigma}{d y }$ [42]	7	0.53	$\chi_{\text{exp+th}}^2$	3.83	3.32	3.33	3.93	3.43	2.24	0.17	1.95
			χ_{exp}^2	7.90	8.42	8.38	8.77	7.24	3.49	0.18	3.19
ATLAS 8 TeV $Z \frac{d\sigma}{d y }$ [68]	60	0.18	$\chi_{\text{exp+th}}^2$	1.08	1.05	1.01	1.09	1.08	1.06	1.02	1.01
			χ_{exp}^2	1.23	1.18	1.11	1.25	1.24	1.28	1.41	1.17
ATLAS 8 TeV $Z \frac{d\sigma}{d y }$ (at Z -peak) [68]	24	0.29	$\chi_{\text{exp+th}}^2$	1.30	1.27	1.27	1.29	1.31	1.30	1.28	1.28
			χ_{exp}^2	1.31	1.28	1.27	1.30	1.31	1.31	1.28	1.31

Table 5. Same as table 4 for the baseline NNPDF4.0 PDF set and for the additional NNPDF4.0-like PDF sets described in the text. Values are displayed separately for the measurement from [42], for the measurement from [68], and for the subset of the latter corresponding to the invariant mass bin of the Z peak.

therefore unlikely that theoretical inaccuracy is a limitation in the description of the ATLAS measurement of [42]. The ATLAS dataset of [68] is described with comparable quality by a PDF set determined from a fit without it (fit (a)); in this case, the description of the ATLAS dataset of [42] does not improve in a significant way. If instead one tries to fit the ATLAS measurement of [42] (fit (b)), the value of $\chi_{\text{exp+th}}^2$ (χ_{exp}^2) improves by about 2σ (8σ). At the same time, the description of the ATLAS measurement of [68] does not change in a significant way. We therefore conclude that the old and new measurements are not in tension between each other. The picture can be further improved if one repeats fit (b) with inclusion of MHOU (fit (d)): in this case, the values of $\chi_{\text{exp+th}}^2$ and χ_{exp}^2 reduce by about another half of a sigma. One may finally wonder whether the ATLAS measurement of [42] is in tension with other datasets included in NNPDF4.0. In this respect, fit (c) reveals that an overly good description of the dataset can be achieved if it is overweighted. The description of the ATLAS measurement of [68] is not significantly altered, in comparison to the other fits, thus confirming that the two measurements are consistent with each other. However, the global fit quality deteriorates significantly: the total χ_{exp}^2 per point increases from 1.16 (in the default NNPDF4.0 fit) to 1.24. Because there are about 4600 fitted data points, this corresponds to a worsening of about 4σ . Inspection of individual dataset figures reveals that this is due to a deterioration in the description of several Drell-Yan measurements, which see the following increase in the value of χ_{exp}^2 : D0 W muon asymmetry production [131] from 1.91 to 4.27 ($n_{\text{dat}} = 9$); ATLAS W , Z production, 7 TeV [132] from 1.67 to 3.07 ($n_{\text{dat}} = 53$); LHCb Z production, 7 TeV [133] from 1.65 to 2.48 ($n_{\text{dat}} = 9$); LHCb W^\pm production, 7 TeV [134] from 1.97 to 4.12 ($n_{\text{dat}} = 29$); and LHCb Z production, 8 TeV [135] from 1.33 to 2.32 ($n_{\text{dat}} = 17$). We therefore conclude that the ATLAS measurement of [42] is in tension with these other datasets.

In summary, the ATLAS measurement of [42] is consistent with its earlier version [68], but in tension with other Drell-Yan measurements included in NNPDF4.0. An acceptable description of it can be achieved if this is included in the fit and if MHOU are taken into account. The value of $\chi_{\text{exp+th}}^2 = 1.95$ is indeed only slightly less than 2σ away from the unit

expectation. This level of disagreement does not appear to be more pathological than that observed for few other datasets included in NNPDF4.0, and is also similar to that observed for other datasets examined in the following. Note however that, for all of the other datasets examined in this work, we will look at the data-theory (dis)agreement only *before* inclusion in a fit. Refitting could possibly improve the overall picture, however investigations along this direction are beyond the scope of this work, as they will be part of the dataset selection involved with a future NNPDF release.

4.3 Top-quark pair production measurements

We continue by discussing the LHC top-quark pair production measurements outlined in section 2, see also table 2. The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$, computed for each of the PDF sets summarised in section 3.3, are reported in table 6. The experimental covariance matrix is regularised as explained in section 3.2 for the following datasets: the ATLAS all-hadronic absolute single-differential distribution in the invariant mass of the top-quark pair and double-differential distribution in the invariant mass and absolute rapidity of the top-quark pair; the ATLAS lepton+jets normalised single-differential distributions in the invariant mass of the top-quark pair and in the transverse momentum of the top quark; and the CMS lepton+jets normalised double-differential distribution in the invariant mass and absolute rapidity of the top-quark pair. See appendix A for the unregularised values of $\chi_{\text{exp+th}}^2$. The breakdown of $\chi_{\text{exp+th}}^2$ into $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 is displayed in figure 5, albeit only for a representative subset of distributions, specifically: the ATLAS lepton+jets normalised cross sections, single-differential in the transverse momentum of the top quark, p_T^t , and in the invariant mass of the top-quark pair, $m_{t\bar{t}}$; the CMS lepton+jets normalised cross sections, single-differential in the absolute rapidity of the top quark and of the top-quark pair, $|y_t|$ and $|y_{t\bar{t}}|$; the ATLAS all-hadronic absolute cross section, double-differential in the invariant mass and absolute rapidity of the top-quark pair; and the CMS lepton+jets normalised cross section, double-differential in the invariant mass and absolute rapidity of the top-quark pair. Histogram plots for the other datasets are collected in figure 15 of appendix B. The data-theory comparison is displayed in figure 6 for the same representative subset of top-quark pair measurements of figure 5. In the case of the ATLAS and CMS double-differential distributions, only the bin at the lowest invariant mass is shown. Additional results are collected in figures 16–17 of appendix B. Note that, for normalised distributions, we consistently do not display the last bin, which is linearly dependent from the others by construction. Hence the number of data points displayed is one unit less than the number of data points reported in table 2.

From inspection of table 6 and of figure 5, we make considerations very similar to those made for Drell-Yan weak boson production measurements at the LHC 13 TeV. Namely, that the values of $\chi_{\text{exp+th}}^2$, computed with different input PDFs, are closer to each other than the corresponding values of χ_{exp}^2 , and that the former are generally rather similar across PDF sets. The only partial exceptions are given by the ATLAS all-hadronic double differential measurement, for which the values of $\chi_{\text{exp+th}}^2$ are away from one by slightly more than 1σ for the MSHT20 and CT18Z sets, and the ATLAS leptons+jet single-differential measurement in the absolute value of the top-quark rapidity, for which the total $\chi_{\text{exp+th}}^2$ computed with NNPDF3.1 is slightly more than 1σ away from one. As far as CMS is concerned, the

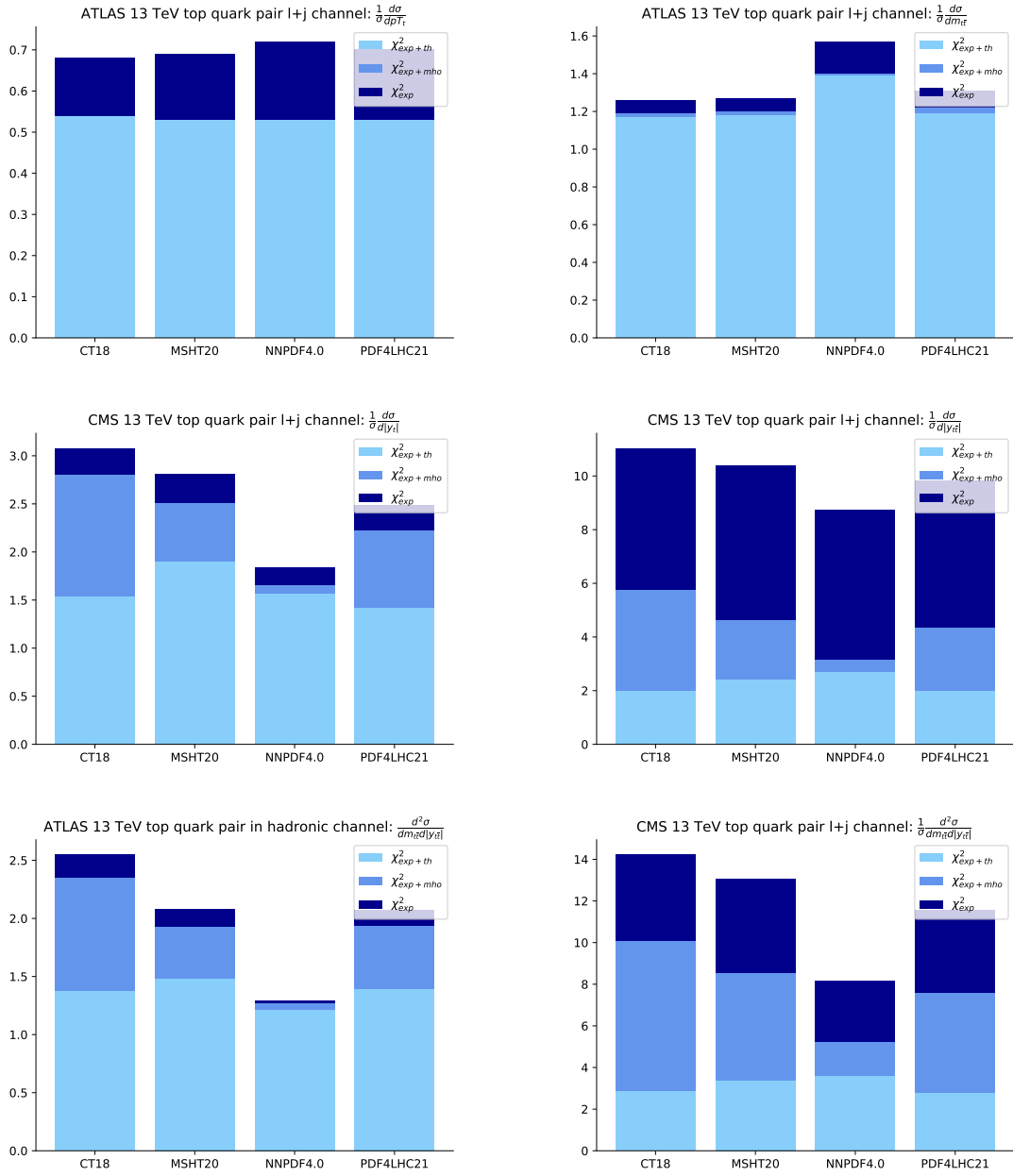


Figure 5. Same as figure 2 for a representative subset of top-quark pair production measurements at the LHC 13 TeV. Specifically, from top to bottom, left to right: the ATLAS lepton+jets normalised cross sections single-differential in the transverse momentum of the top quark p_T^t ($n_{\text{dat}} = 8$, $\sqrt{2/n_{\text{dat}}} = 0.50$), and in the invariant mass of the top-quark pair $m_{t\bar{t}}$ ($n_{\text{dat}} = 9$, $\sqrt{2/n_{\text{dat}}} = 0.47$); the CMS lepton+jets normalised cross sections, single-differential in the absolute rapidity of the top quark and of the top-quark pair $|y_t|$ ($n_{\text{dat}} = 11$, $\sqrt{2/n_{\text{dat}}} = 0.43$) and $|y_{t\bar{t}}|$ ($n_{\text{dat}} = 10$, $\sqrt{2/n_{\text{dat}}} = 0.45$), the ATLAS all-hadronic absolute cross section, double-differential in the invariant mass and absolute rapidity of the top-quark pair ($n_{\text{dat}} = 11$, $\sqrt{2/n_{\text{dat}}} = 0.43$); and the CMS lepton+jets normalised cross section, double-differential in the invariant mass and absolute rapidity of the top-quark pair ($n_{\text{dat}} = 35$, $\sqrt{2/n_{\text{dat}}} = 0.24$). Histogram plots for the other datasets are collected in figure 15 of appendix B.

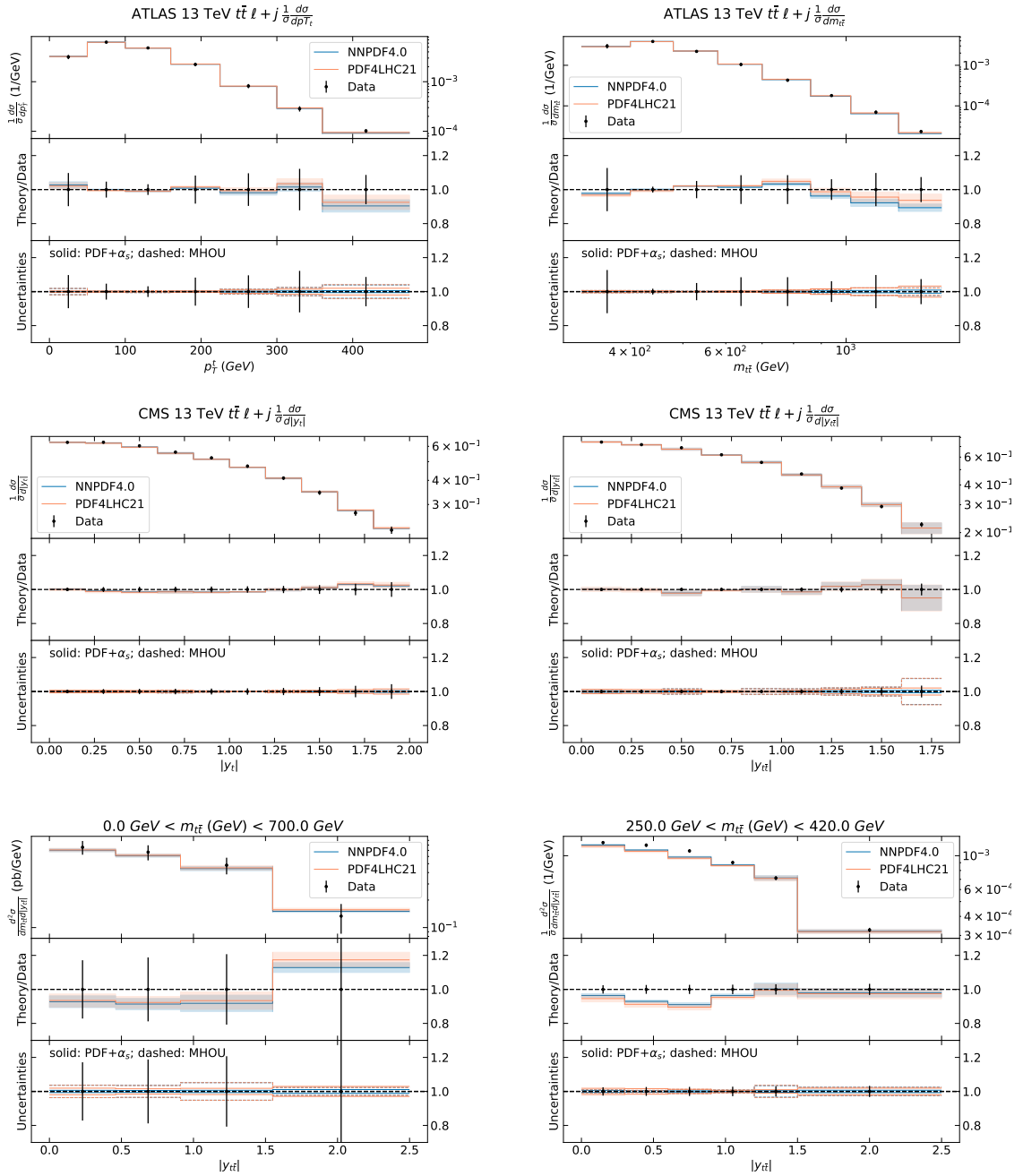


Figure 6. Same as figure 3 for the same representative subset of top-quark pair measurements of figure 5. In the case of the ATLAS (bottom left) and CMS (bottom right) double-differential distributions, only the bin at the lowest invariant mass is shown. Additional results are collected in figures 16–17 of appendix B.

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d\sigma}{dm_{t\bar{t}}}$	9	0.47	$\chi_{\text{exp+th}}^2$	0.84	0.99	0.97	0.94	0.97	0.86	0.81	0.96	0.93
			χ_{exp}^2	0.88	1.21	1.16	1.15	1.12	0.91	0.84	1.13	1.06
ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{1}{\sigma} \frac{d\sigma}{d y_{t\bar{t}} }$	12	0.41	$\chi_{\text{exp+th}}^2$	0.62	0.78	0.77	0.85	0.74	0.64	0.68	0.73	0.73
			χ_{exp}^2	0.68	0.85	0.83	0.95	0.79	0.67	0.71	0.82	0.78
ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	11	0.43	$\chi_{\text{exp+th}}^2$	0.92	1.38	1.39	1.42	1.48	1.12	1.22	1.22	1.39
			χ_{exp}^2	1.05	2.55	2.38	2.84	2.08	1.20	1.29	2.11	2.07
ATLAS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{dm_{t\bar{t}}}$	9	0.47	$\chi_{\text{exp+th}}^2$	1.41	1.17	1.17	1.04	1.18	1.46	1.39	1.20	1.19
			χ_{exp}^2	1.67	1.26	1.26	1.12	1.27	1.65	1.57	1.32	1.31
ATLAS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{dp_T^t}$	8	0.50	$\chi_{\text{exp+th}}^2$	0.56	0.54	0.54	0.52	0.53	0.56	0.53	0.53	0.53
			χ_{exp}^2	0.76	0.68	0.68	0.67	0.69	0.77	0.72	0.68	0.70
ATLAS 13 TeV $t\bar{t} \ell + j$ $\frac{d\sigma}{d y_{t\bar{t}} }$	5	0.63	$\chi_{\text{exp+th}}^2$	1.39	1.05	1.09	0.92	1.10	1.70	1.58	1.09	1.15
			χ_{exp}^2	1.62	1.17	1.19	1.00	1.14	1.86	1.62	1.26	1.29
ATLAS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{d y_{t\bar{t}} }$	7	0.53	$\chi_{\text{exp+th}}^2$	0.57	0.43	0.42	0.58	0.47	0.58	0.42	0.42	0.39
			χ_{exp}^2	0.74	0.57	0.55	0.99	0.66	0.65	0.47	0.56	0.47
CMS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{dm_{t\bar{t}}}$	15	0.37	$\chi_{\text{exp+th}}^2$	0.24	0.49	0.51	0.53	0.57	0.29	0.33	0.42	0.44
			χ_{exp}^2	0.37	1.38	1.30	1.44	1.15	0.39	0.42	1.14	0.98
CMS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	35	0.24	$\chi_{\text{exp+th}}^2$	2.77	2.89	2.87	2.76	3.36	3.01	3.61	2.81	2.81
			χ_{exp}^2	8.37	14.2	13.7	16.6	13.1	7.31	8.14	13.1	11.6
CMS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{dp_T^t}$	16	0.35	$\chi_{\text{exp+th}}^2$	0.78	0.62	0.63	0.66	0.64	0.79	0.81	0.63	0.65
			χ_{exp}^2	1.31	0.68	0.70	0.69	0.72	1.24	1.17	0.74	0.78
CMS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{d y_{t\bar{t}} }$	11	0.43	$\chi_{\text{exp+th}}^2$	1.07	1.54	1.57	1.81	1.90	1.22	1.57	1.38	1.42
			χ_{exp}^2	1.61	3.08	2.94	4.02	2.81	1.46	1.84	2.77	2.49
CMS 13 TeV $t\bar{t} \ell + j$ $\frac{1}{\sigma} \frac{d\sigma}{d y_{t\bar{t}} }$	10	0.45	$\chi_{\text{exp+th}}^2$	0.94	2.01	1.89	2.16	2.44	1.76	2.71	1.53	2.00
			χ_{exp}^2	8.65	11.0	10.7	12.9	10.4	8.06	8.72	10.6	9.82

Table 6. Same as table 3 for the ATLAS and CMS top-quark pair production measurements at the LHC 13 TeV.

leptons+jet double differential distribution is poorly described by all PDF sets, with the total $\chi_{\text{exp+th}}^2$ being away from one by more than 3σ for all PDF sets. Less dramatic, but still significant, is the spread of $\chi_{\text{exp+th}}^2$ values across PDF sets for the CMS lepton+jets single-differential distribution in the absolute rapidity of the top-quark and of the top-quark pair: in the former case, predictions obtained with CT18Z and MSHT20 are 2σ away from one, whereas predictions obtained with MSHT20 and NNPDF4.0 are about 3σ away from one (with NNPDF4.0 performing a little worse than MSHT20 by a quarter of a sigma).

The breakdown of $\chi_{\text{exp+th}}^2$ into its theoretical components depends on the dataset. The component due to MHO, which is relatively independent from the PDF set, prevails over the PDF+ α_s component in the ATLAS lepton+jets distributions differential in the transverse

momentum of the top quark and in the invariant mass of the top-quark pair, and in the CMS lepton+jets distribution differential in the absolute rapidity of the top-quark pair. The PDF+ α_s component prevails in the other datasets, although it depends on the PDF set: it is generally larger for the CT18 and PDF4LHC21 PDF sets, which are affected by the largest uncertainties, see figure 1, whereas it is almost immaterial for NNPDF4.0, which has the smallest PDF uncertainties.

Overall, the quality of the data description is generally good, being $\chi_{\text{exp+th}}^2 \sim 1$ for all the datasets, except in the case of the CMS normalised single-differential distribution in the absolute rapidity of the top-quark pair, and double-differential distribution in the absolute rapidity and invariant mass of the top-quark pair, for which $\chi_{\text{exp+th}}^2 \sim 2 - 3$. Discrepancies between data and theory that may lead to these results are seen in figure 6, where experimental data and theoretical predictions are generally well aligned to each other, within their uncertainties, except, precisely, for the aforementioned datasets. Understanding the reason for this behaviour, which is common to most PDF sets, is left to future investigations.

4.4 Single-inclusive jet and di-jet production measurements at the LHC

We now turn to LHC single-inclusive jet and di-jet production measurements outlined in section 2. The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$ are reported in table 7. The experimental covariance matrix is regularised as explained in section 3.2 for all the datasets. Without regularisation, the values of χ_{exp}^2 are very poor, as they are away from one by more than 10-20 σ , independently of the input PDF set. See appendix A for the unregularised values of $\chi_{\text{exp+th}}^2$. Clearly the results that we present here do depend on regularisation. However, as discussed in appendix A, this dependence does not affect our ability to discriminate how well different PDF sets describe the data, which is the goal of this work. If, instead, we were interested to characterise the datasets for inclusion in a PDF determination (or not), we would consider other ways of decorrelating uncertainties, for instance by identifying uncertainties that, for experimental reasons, are more likely to be overcorrelated, see e.g. [111]. In this sense, the main message conveyed by the numbers in table 7 is that the single-inclusive jet and di-jet datasets require additional investigations on the understanding of uncertainty correlations. Only after accomplishing these investigations, which are beyond the scope of this work, one may be able to better judge how the various PDF sets comparatively generalise on them.

The breakdown of $\chi_{\text{exp+th}}^2$ into $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 after regularisation is displayed in figure 7. The data-theory comparison is displayed in figures 8, 9, and 10, respectively for the ATLAS and CMS single-inclusive jet, and for the ATLAS di-jet measurements. In the first and second (third) cases, we plot the double differential cross section as a function of the transverse momentum of the leading jet, p_T^j (the invariant mass of the two jets, m_{jj}), for the two outermost bins of the absolute value of the jet rapidity, $|y_j|$ (of the two-jet rapidity separation $|y^*|$). The other bins are displayed, respectively, in figures 18, 19, and 20 of appendix B.

From inspection of table 7 and of figure 7, very similar remarks can be drawn for the three considered datasets. First, when theory errors are not included in the computation of the χ^2 , the NNPDF4.0 PDF set performs better than any of the others, in the sense that the NNPDF4.0 χ_{exp}^2 is the closest to unity among all PDF sets, although the χ_{exp}^2 is still away

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
ATLAS 13 TeV incl. jet $\frac{d^2\sigma}{dp_T d y }$	177	0.11	$\chi^2_{\text{exp+th}}$	1.85	1.56	1.64	1.38	1.67	1.21	1.51	1.20	1.25
			χ^2_{exp}	2.32	2.48	2.47	2.50	2.53	2.98	1.95	3.02	2.40
CMS 13 TeV incl. jet $R_{0.4} \frac{d^2\sigma}{dp_T d y }$	78	0.16	$\chi^2_{\text{exp+th}}$	1.64	1.58	1.60	1.52	1.64	1.47	1.50	1.48	1.43
			χ^2_{exp}	2.05	2.29	2.25	2.26	2.23	2.21	2.02	2.30	2.18
ATLAS 13 TeV di-jets $\frac{d^2\sigma}{dm_{jj} d y^* }$	136	0.12	$\chi^2_{\text{exp+th}}$	1.13	1.08	1.09	1.05	1.16	1.09	1.15	1.01	0.96
			χ^2_{exp}	1.25	1.49	1.47	1.48	1.41	1.37	1.29	1.42	1.41

Table 7. Same as table 3 for the ATLAS and CMS single-inclusive and di-jet production measurements at the LHC 13 TeV.

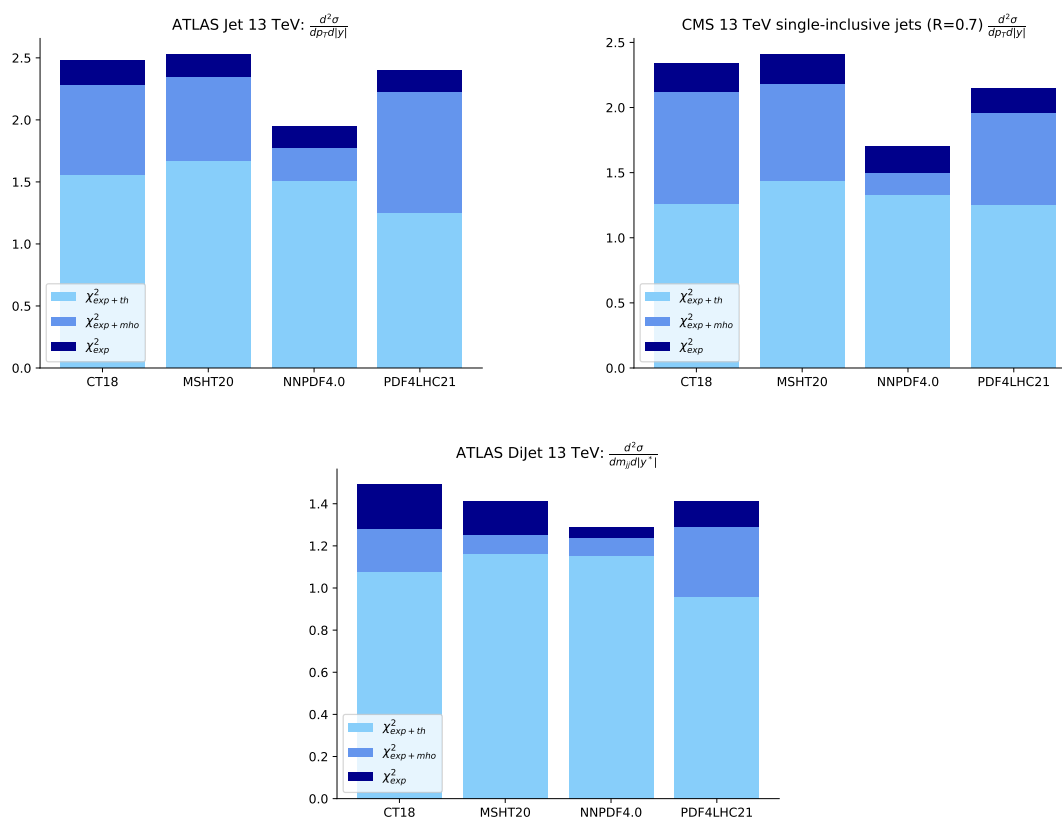


Figure 7. Same as figure 2 for the ATLAS and CMS single-inclusive (with $n_{\text{dat}} = 177$ and $n_{\text{dat}} = 78$ respectively, and $\sqrt{2/n_{\text{dat}}} = 0.11$ and 0.16 correspondingly) and di-jet production measurements at the LHC 13 TeV ($n_{\text{dat}} = 136$, $\sqrt{2/n_{\text{dat}}} = 0.12$).

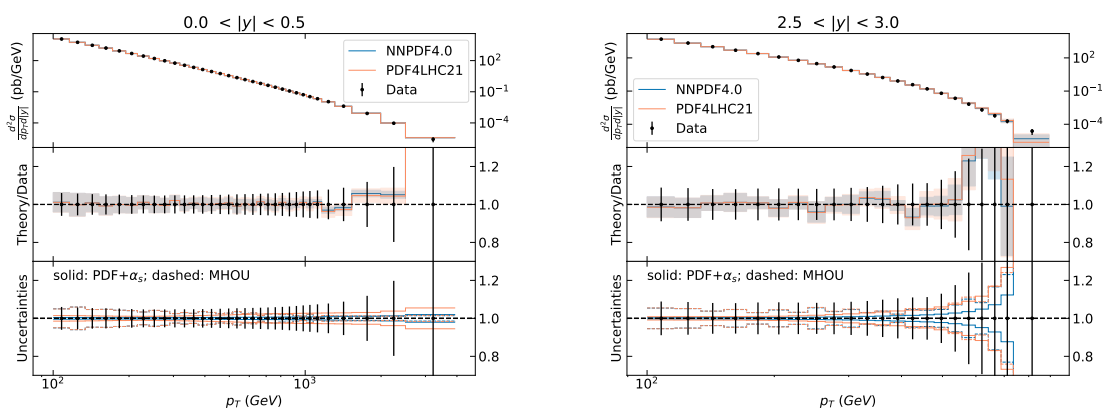


Figure 8. Same as figure 3 for the ATLAS single-inclusive jet double differential cross section as a function of the transverse momentum of the leading jet, p_T^j , for the two outermost bins of the absolute value of the jet rapidity, $|y_j|$. The other bins are displayed in figure 18 of appendix B.

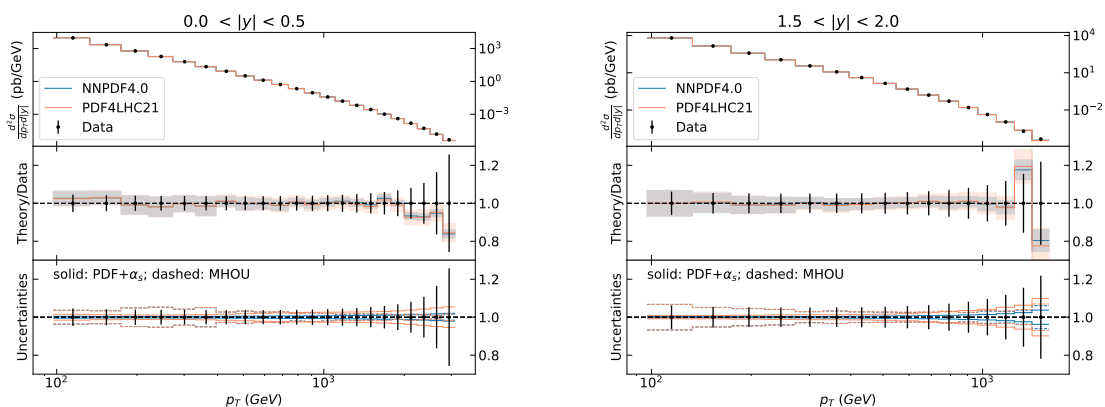


Figure 9. Same as figure 8 for the CMS single-inclusive jet double differential cross section. The other bins are displayed in figure 19 of appendix B.

from one by more than 10σ . Some PDF sets may lead to comparatively larger values of χ_{exp}^2 , such as APMP16, however the statistical significance of these fluctuations must be seen in units of the χ^2 standard deviation, as we will further discuss in section 4.6. Second, once all the theory errors are included, the values of $\chi_{\text{exp+th}}^2$ become relatively close, irrespective of the input PDF set used for their computation. However, in the case of single-inclusive jets, they are all a few units away from one, suggesting that once these datasets are included in the fit, they might have a significant pull on the gluon PDF. The values of $\chi_{\text{exp+th}}^2$ being relatively close suggests also that, except perhaps for ABMP16, which continues to display rather large values of $\chi_{\text{exp+th}}^2$ even after inclusion of theoretical uncertainties, it may be difficult to discriminate the quality of the predicting power of the various PDF sets based solely on these measurements. Third, the relatively homogeneous values of $\chi_{\text{exp+th}}^2$ occur despite the input PDF sets have very different uncertainties. For instance, PDF4LHC21 uncertainties are twice the NNPDF4.0 uncertainties, see figure 1. The breakdown of the theoretical uncertainty into

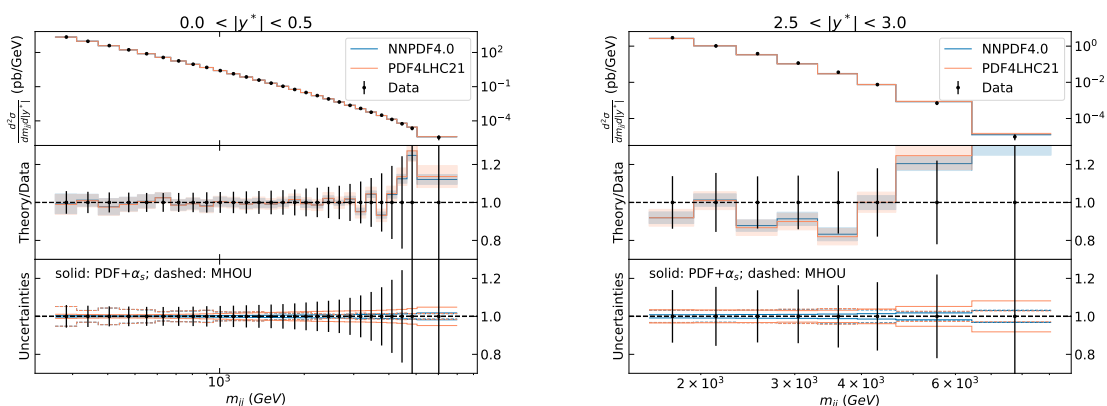


Figure 10. Same as figure 3 for the ATLAS di-jet double differential cross section, as a function of the invariant mass of the di-jet pair, m_{jj} , for the two outermost bin in the absolute rapidity separation between the two jets. The other bins are displayed in figure 20 of appendix B.

its various components can be different depending on the PDF set. The MHO uncertainty remains more or less the same for all PDF sets. Conversely, the PDF+ α_s uncertainty is the smallest for NNPDF4.0. This is consistent with the fact that NNPDF4.0 PDF uncertainties are typically the smallest among all the PDF sets considered, see figure 1. Finally, it is interesting to observe that the balance between the various components of the theoretical uncertainty depend on the kinematics. From figures 8–10, we see that the PDF+ α_s (MHO) uncertainty dominates at small (large) p_T^j or m_{jj} .

4.5 Single-inclusive jet and di-jet production measurements at HERA

We finally discuss the HERA single-inclusive jet and di-jet production measurements outlined in section 2.3. The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$ are reported in table 8. The experimental covariance matrix of the H1 low- Q^2 single-inclusive jet and di-jet measurements is regularised as explained in section 3.2. The unregularised values of $\chi_{\text{exp+th}}^2$ are reported in appendix A. The breakdown of $\chi_{\text{exp+th}}^2$ into $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 is displayed in figure 11, albeit only for the H1 data. The data-theory comparison is displayed in figure 12 for the highest Q^2 bin of the H1 single-inclusive jet and di-jet differential cross sections as a function, respectively, of the transverse momentum of the leading jet and of the average transverse momentum of the jet pair. Histograms plots for the ZEUS measurements and data-theory comparison plots for the remaining H1 bins and for all of the ZEUS bins are collected in figures 21–28 of appendix B.

From inspection of table 8 and figure 11, we observe that the values of $\chi_{\text{exp+th}}^2$ and of χ_{exp}^2 are very similar when different input PDF sets are used. All PDF sets generalise equally well on these datasets. The largest component of $\chi_{\text{exp+th}}^2$ is due to MHO, in a proportion which is roughly the same across PDF sets. The PDF+ α_s component of $\chi_{\text{exp+th}}^2$ is almost immaterial (for the H1 low- Q^2 dataset), very small (for the H1 high- Q^2 single-inclusive jet dataset), or as large as the MHO component (for the H1 high- Q^2 di-jet dataset). The quality of the data description is generally very good, with $\chi_{\text{exp+th}}^2 \sim 1$ for all the datasets, except for the H1 high- Q^2 dataset, in which case $\chi_{\text{exp+th}}^2 \sim 2$. Investigations into the reasons

Dataset	n_{dat}	$\sqrt{2/n_{\text{dat}}}$		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
H1 incl. jet (low Q^2) $\frac{d^2\sigma}{dQ^2 dp_T}$	37	0.23	$\chi_{\text{exp+th}}^2$	1.64	1.61	1.61	1.67	1.61	1.70	1.74	1.61	1.73
			χ_{exp}^2	7.68	2.17	2.14	2.11	2.16	2.16	2.12	2.17	2.14
H1 incl. jet (high Q^2) $\frac{d^2\sigma}{dQ^2 dp_T}$	24	0.29	$\chi_{\text{exp+th}}^2$	1.62	1.66	1.62	1.63	1.64	1.49	1.63	1.58	1.59
			χ_{exp}^2	2.40	2.28	2.20	2.18	2.27	2.43	2.42	2.33	2.27
ZEUS incl. jet (low lumi.) $\frac{d^2\sigma}{dQ^2 dE_T}$	30	0.26	$\chi_{\text{exp+th}}^2$	0.67	0.69	0.68	0.67	0.68	0.66	0.65	0.68	0.67
			χ_{exp}^2	0.69	0.71	0.70	0.69	0.70	0.69	0.67	0.70	0.69
ZEUS incl. jet (high lumi.) $\frac{d^2\sigma}{dQ^2 dE_T}$	30	0.26	$\chi_{\text{exp+th}}^2$	0.77	0.77	0.77	0.76	0.78	0.77	0.76	0.77	0.77
			χ_{exp}^2	0.82	0.83	0.82	0.80	0.82	0.84	0.81	0.83	0.82
H1 di-jets (low Q^2) $\frac{d^2\sigma}{dQ^2 d(p_T)}$	37	0.23	$\chi_{\text{exp+th}}^2$	1.37	1.39	1.38	1.37	1.39	1.42	1.44	1.36	1.44
			χ_{exp}^2	11.0	1.75	1.73	1.68	1.75	1.82	1.78	1.77	1.75
H1 di-jets (high Q^2) $\frac{d^2\sigma}{dQ^2 d(p_T)}$	24	0.29	$\chi_{\text{exp+th}}^2$	2.21	2.03	2.00	1.95	2.03	1.84	2.12	1.94	1.97
			χ_{exp}^2	2.63	2.47	2.37	2.32	2.42	2.65	2.63	2.51	2.45
ZEUS di-jets $\frac{d^2\sigma}{dQ^2 d(E_T)}$	22	0.30	$\chi_{\text{exp+th}}^2$	0.81	0.75	0.75	0.71	0.78	0.90	0.83	0.77	0.79
			χ_{exp}^2	1.49	1.29	1.27	1.24	1.32	1.71	1.63	1.37	1.42

Table 8. Same as table 3 for HERA single-inclusive jet and di-jet data.

for this behaviour, which is consistent throughout PDF sets, will be left to future work. For now, we remark that the agreement between experimental data and the corresponding theoretical predictions, as seen in figure 12, is generally good, except for specific bins that display larger fluctuations between the two.

4.6 Combined interpretation

We now combine the results described in the previous sections to gather the overall agreement between the considered experimental data and the corresponding theoretical predictions. To this purpose, in figure 13, we display $\Delta\chi^{2(i)}$, the relative change in the total $\chi_{\text{exp+th}}^2$ due to the change of input PDF set with respect to the average $\chi_{\text{exp+th}}^2$ over PDF sets, see eq. (3.16). The PDF sets considered here are ABMP16, CT18, MSHT20, NNPDF4.0, and PDF4LHC21. All the datasets listed in table 1 are considered, except for the 8 TeV ATLAS Drell-Yan rapidity distribution [42]. The reason being that this dataset, extensively discussed in section 4.2, is included in MSHT20 and NNPDF4.0 in the form of an earlier analysis [68], whereas all the other datasets are not included in any PDF set. Furthermore, all the other data sets are for the LHC Run II. The datasets are grouped by category: LHC Drell-Yan, LHC top-quark pair, LHC single-inclusive jet and di-jet, and HERA single-inclusive jet and di-jet production cross sections. The circumference corresponding to $\Delta\chi^2 = 0$ is highlighted with a solid curve. In figure 14 we display, in the same format, $\Delta n_{\sigma}^{(i)}$, the difference between the total $\chi_{\text{exp+th}}^2$ computed with the i -th PDF set and the average $\chi_{\text{exp+th}}^2$ over PDF sets, normalised to the standard deviation of the χ^2 distribution, see eq. (3.18). Figures 13 and 14 should be inspected together: the latter provides an assessment of the statistical significance

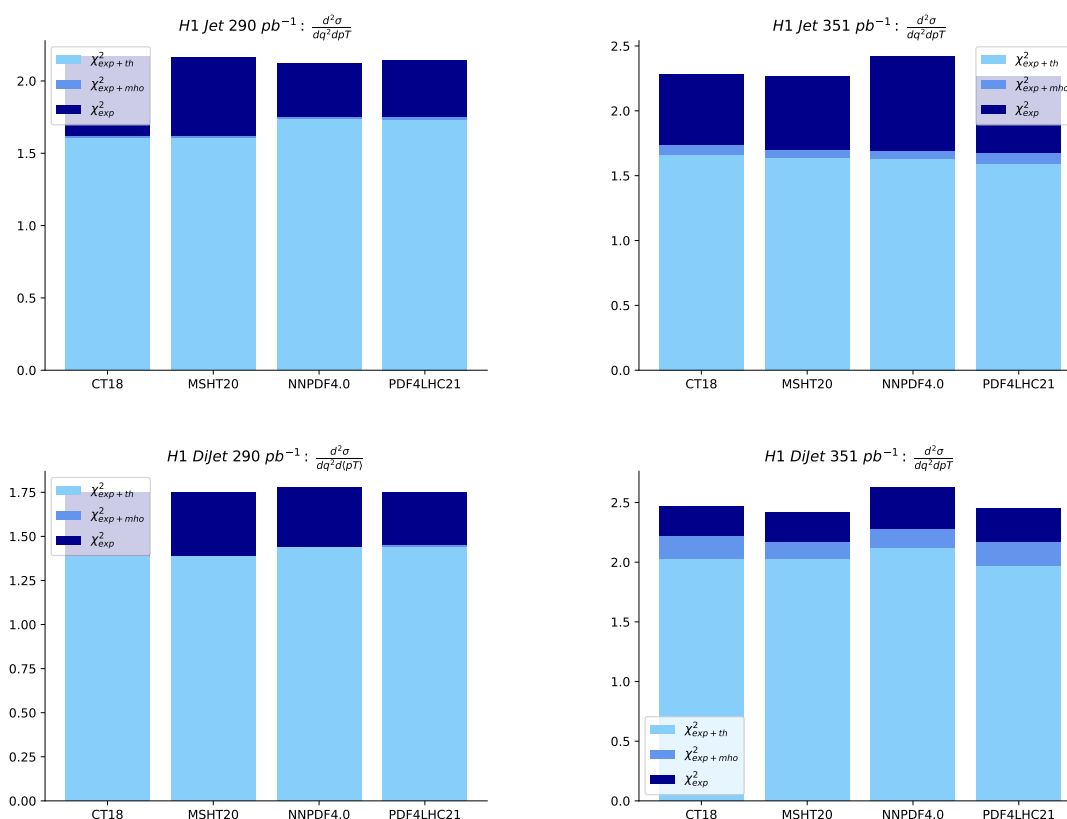


Figure 11. Same as figure 2 for the H1 single-inclusive jet (top) and di-jet (bottom) datasets. For the measurements on the left plots $n_{\text{dat}} = 37$ and $\sqrt{2/n_{\text{dat}}} = 0.23$, while for the measurements on the right plots $n_{\text{dat}} = 24$ and $\sqrt{2/n_{\text{dat}}} = 0.29$.

of fluctuations from the average $\Delta\chi^2 = 0$ seen in the former, in units of the χ^2 standard deviation. Large fluctuations may have low statistical significance if a dataset has a small number of data points and the other way around.

On the basis of figures 13 and 14, we conclude that the various classes of datasets are described to a different level of accuracy. However, whereas the value of $\Delta\chi^2$ displays sizeable fluctuations depending on the input PDF set, especially in the top-quark pair and jet sectors, we realise that discrepancies with respect to the average over PDF sets is almost always within $\Delta n_\sigma = 1$. The most relevant excess occurs with the ABMP16 PDF set in the case of the ATLAS and CMS single-inclusive jet measurements, and with the NNPDF4.0 PDF set in the case of the CMS double-differential and rapidity-differential top-quark pair measurements. In these cases, the excess is between one and three sigma. This fact may be explained by assuming that these measurements disfavor the softer (harder) large- x gluon of ABMP16 (NNPDF4.0). We also note an anomalous deficiency, close $\Delta n_\sigma = 3$, with the PDF4LHC21 PDF set in the case of the ATLAS single-inclusive jet measurements, and with the ABMP16 PDF set in the case of the CMS rapidity-differential top-quark pair measurement. We therefore conclude that, whereas HERA jet and LHC Drell-Yan measurements may not be able to discriminate between PDF sets, LHC jet and top-quark pair measurements may help

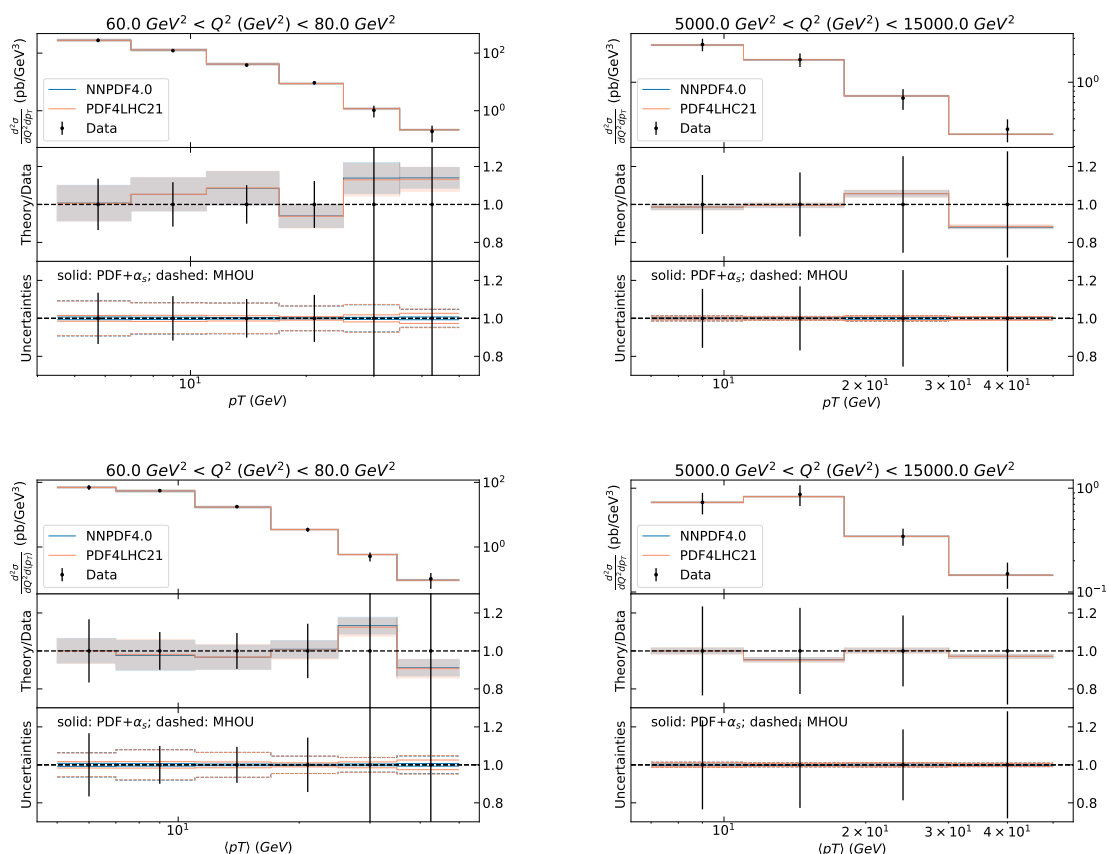


Figure 12. Same as figure 3 for the largest Q^2 bins of the H1 single-inclusive jet (top) and di-jet (bottom) datasets. All the other bins are collected in figures 22–25 of appendix B.

put stronger constraints on PDFs, especially those datasets for which the largest fluctuations among different PDF sets are observed in terms of Δn_σ .

As for the general trend displayed by individual PDF sets, on top of the aforementioned dataset-specific considerations, we remark two interesting facts. First, the NNPDF4.0 PDF set, despite displaying the smallest uncertainties among all the PDF sets considered in this work (see figure 1), provides a description of the data which is overall not worse than the others (with the aforementioned few exceptions). We therefore conclude that theoretical predictions obtained with NNPDF4.0 are overall as accurate as those obtained with the other PDF sets, despite their smaller PDF uncertainties, once experimental, MHO, and α_s uncertainties are taken into account, and the regularisation of the covariance matrix is applied when needed. This is true on average on the examined datasets: as discussed, there are cases in which NNPDF4.0 performs better than other PDF sets, others in which it performs worse than these, and others in which the experimental precision of the datasets cannot discriminate among different PDF sets. Second, the PDF4LHC21 PDF set generally displays the value of $\Delta\chi^2$ and Δn_σ closest to zero among all the PDF sets considered in this work. This fact is unsurprising, given that PDF4LHC21 is the unweighted average of the CT18, MSHT20 and NNPDF3.1 PDF sets. Deviations from the mean $\Delta\chi^2 = 0$ and $\Delta n_\sigma = 0$, obtained with

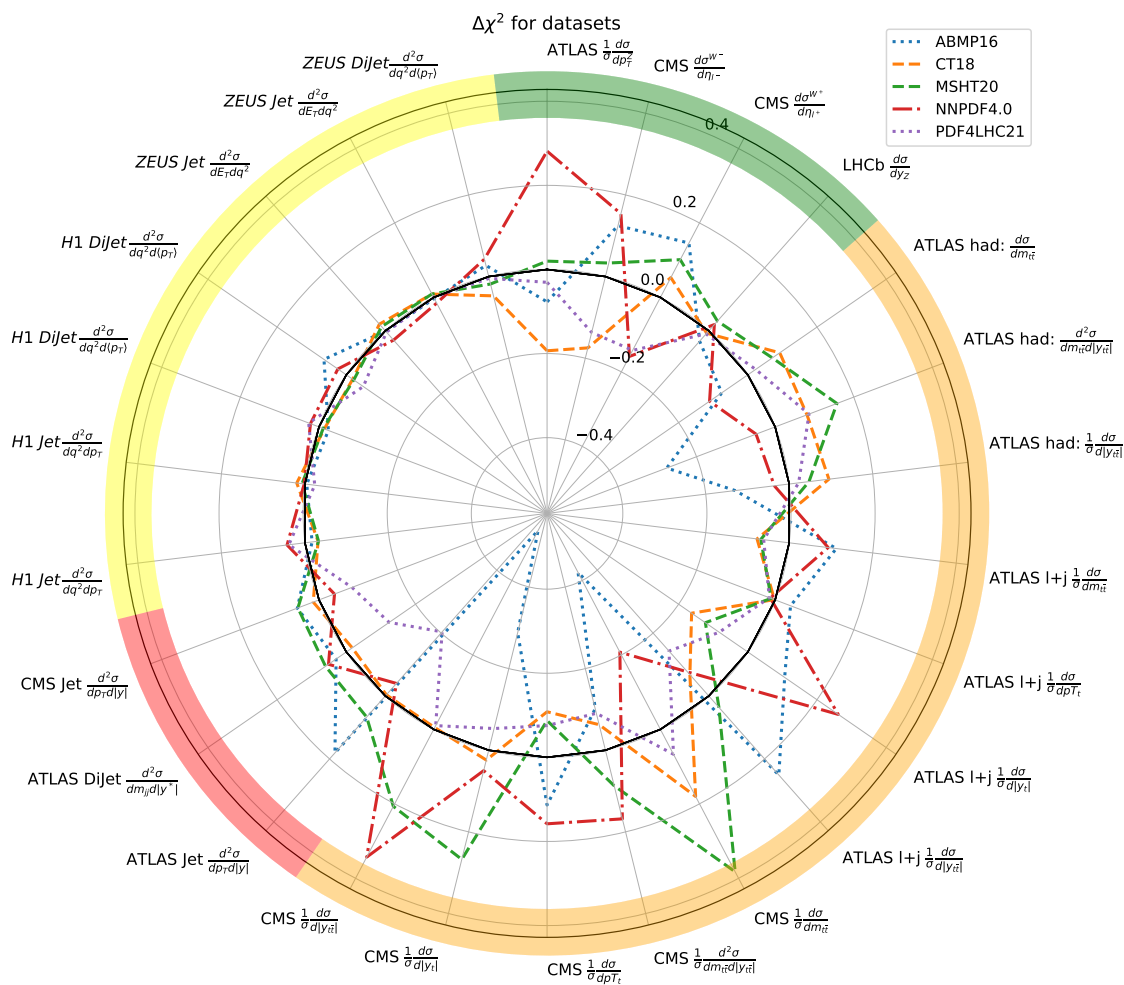


Figure 13. The relative change in the total $\chi^2_{\text{exp+th}}$ due to the change of i -th input PDF set, $\Delta\chi^{2(i)}$, with respect to the average $\chi^2_{\text{exp+th}}$ over PDF sets, see eq. (3.16). The PDF sets considered here are ABMP16, CT18, MSHT20, NNPDF4.0, and PDF4LHC21. The datasets are grouped by category: LHC Drell-Yan, LHC top-quark pair, LHC single-inclusive jet and di-jet, and HERA single-inclusive jet and di-jet production cross sections. The circumference corresponding to $\Delta\chi^2 = 0$ is highlighted with a solid curve.

individual PDF sets, cancel out by construction. In this sense, PDF4LHC21 is a conservative PDF set, as already illustrated in [26], although it remains the least precise.

5 Summary and outlook

In this paper we have compared theoretical predictions, computed at NNLO accuracy in perturbative QCD using different input PDF sets, with a wide array of experimental measurements, typically not yet included in PDF determination. Specifically, we have

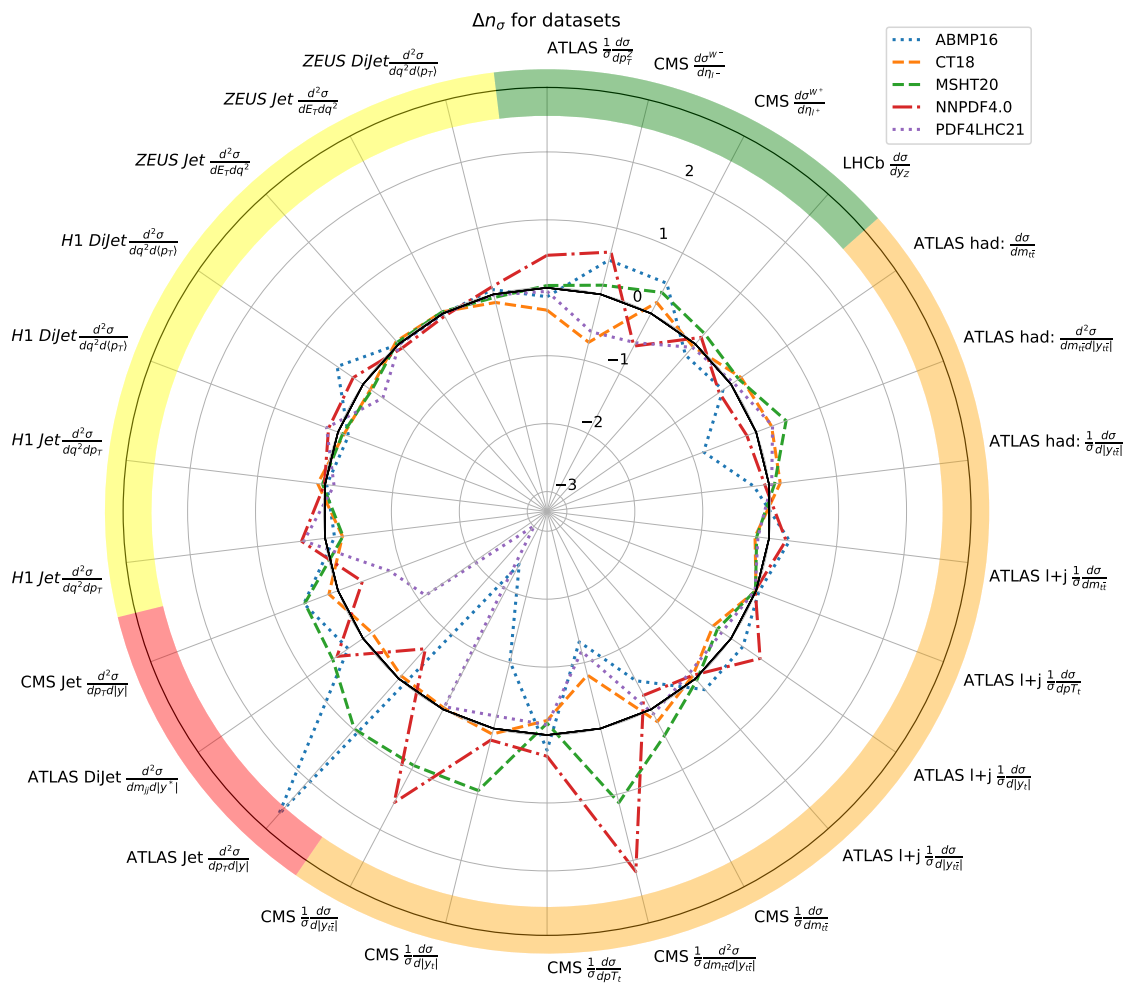


Figure 14. Same as figure 13 now for Δn_σ , see eq. (3.18).

considered differential cross sections measured at the LHC, for Drell-Yan gauge boson, top-quark pair, single-inclusive jet and di-jet production, and at HERA, for single-inclusive jet and di-jet production. We have considered the most widely used PDF sets in LHC experimental analyses, namely, ABMP16, CT18 (and its variants), MSHT20, NNPDF3.1, NNPDF4.0, PDF4LHC15, and PDF4LHC21. We have accounted for all the relevant sources of experimental and theoretical uncertainties, in particular due to PDFs, α_s , and MHOUs.

The aim of our work has been twofold. First, to test the predictive power of different PDF sets, by assessing the goodness with which they describe the datasets not included in their determination. Second, to quantify the various sources of uncertainty that enter theoretical predictions, specifically PDF, α_s , and MHO uncertainties. These two objectives are becoming increasingly relevant given the ever higher precision of LHC experiments to determine SM parameters, such as the strong coupling $\alpha_s(m_Z)$, the W -boson mass m_W , and

the effective lepton mixing angle $\sin^2 \theta_{\text{eff}}^\ell$. This precision is now comparable to, if not better than, that obtained at LEP. This outstanding result requires a careful estimate of all of the sources of uncertainties that accompany it, in particular the PDF uncertainty, which is often dominant in LHC measurements.

In this work we have considered the data-theory agreement between predictions obtained with a broad range of input PDF sets and experimental data, but we have not included any of the examined datasets in a PDF determination. Only the simultaneous inclusion of a given subset of statistically independent datasets considered in this work in future PDF determinations will allow us to determine how the resulting PDFs can adapt to the data, and possibly improve the description of the data, as well as the precision of PDFs.

The main outcome of our investigations is summarised in the overview plots presented in section 4.6. We have found that the CT18, MSHT20, NNPDF4.0, and PDF4LHC21 PDF sets provide a comparable description of all of the datasets considered in this work, once all sources of theoretical uncertainty are taken into account. We have therefore concluded that all PDF sets have a similar predictive power and generalise similarly well to unseen data. Incorporating PDF, α_s and MHO uncertainties is crucial to reach this conclusion. These outcomes may seem counter-intuitive given that individual PDF sets differ among each other for their central values and uncertainties, by an amount that is not always encompassed by the latter. Within this general picture, the NNPDF4.0 and PDF4LHC21 sets represent opposite cases. On the one hand, the NNPDF4.0 set has by far the smallest uncertainties of all PDF sets, hence it is the most precise. In spite of this fact, it describes the examined data, on average, as well as the other PDF sets. On the other hand, the PDF4LHC21 set has some of the largest uncertainties of all PDF sets, hence it is the least precise. This is by construction, given that it is the combination of three different PDF sets. However, the fact that it describes the data as well as the other PDF sets means that it does not need to be as accurate as these.

The only exception to this overall trend is represented by the ATLAS 8 TeV inclusive measurement of the Z rapidity distributions extrapolated to the full phase space, which underlies the recent $\alpha_s(m_Z)$ extraction from the companion measurement differential in the transverse momentum of the Z boson. In this case we have found that, despite the excellent agreement of NNPDF4.0 theoretical predictions with the central values of the experimental data, the peculiar slope in rapidity, combined with the dominance of the luminosity normalisation uncertainty, leads to a poor χ^2 . The χ^2 is instead better for other PDF sets because of their larger PDF uncertainty. This may therefore be a case in which the accuracy of the NNPDF4.0 set does not match its very high precision. We have investigated whether this is truly the case. We have found that using variants of the NNPDF4.0 set that incorporate MHO and/or aN³LO corrections improves the χ^2 only marginally. We have also observed that the NNPDF4.0 set and any of its variants provide an excellent description of the earlier ATLAS measurement of the Z boson rapidity distributions based on the same collider events. We have finally determined that the χ^2 of this dataset can be lowered only if it is included, and overweighted, in a fit, at the price of a slight deterioration in the description of the other datasets. We have therefore concluded that there are residual tensions between this dataset and the other datasets in NNPDF4.0.

Two important by-products of this work have been the computation of fast interpolation grids, accurate to NNLO, and the implementation of the experimental information, in the NNPDF format, for all the considered datasets. These facts will allow us to streamline their inclusion in future NNPDF releases. The fast interpolation framework has two major advantages. First, we will abandon the use of K -factors to account for NNLO corrections in partonic matrix elements. Second, we will be able to readily vary the renormalisation and factorisation scales in the computation of theoretical predictions and determine MHOUs.

To conclude, as the LHC experiments finalise the Run II legacy measurements, start to release datasets based on the Run III luminosity, and prepare for the HL-LHC era, our analysis demonstrates the importance of testing the predictive power of PDFs on a broad set of high-precision measurements with state-of-the-art theoretical predictions, which must crucially include all possible sources of theoretical uncertainty. The methodology laid out in this work can be applied to any upcoming and future LHC measurements that may eventually provide a clear guidance concerning which PDF sets are preferred by the experimental data.

Acknowledgments

We thank our colleagues of the NNPDF Collaboration for discussions on the results presented in this work, in particular C. Schwan for his work on maintaining and supporting the PINEAAPPL library. We thank the authors of NNLOJET and MATRIX, in particular A. Huss, S. Devoto, and S. Kallweit, for providing access to preliminary versions of their codes interfaced to PINEAAPPL. We would like to thank the experimentalists who have helped us interpreting and implementing their measurements, in particular Stefano Camarda, Marco Cipriani, Josh Bendavid, Regina Demina and Otto Hindrichs. E.R.N. is supported by the Italian Ministry of University and Research (MUR) through the “Rita Levi-Montalcini” Program. M.U. and M.N.C. are supported by the European Research Council under the European Union’s Horizon 2020 research and innovation Programme (grant agreement n.950246), and partially by the STFC consolidated grant ST/T000694/1 and ST/X000664/1. J.R. is partially supported by NWO, the Dutch Research Council. J.R. and T.R.R. are supported by an ASDI grant from the Netherlands eScience Center (NLeSC). R.S. is supported by the U.K. Science and Technology Facility Council (STFC) consolidated grants ST/T000600/1 and ST/X000494/1.

A Impact of the experimental covariance matrix regularisation

As discussed in section 3.2, the experimental covariance matrix of all the datasets with $Z_{\mathcal{L}} > 4$ is regularised by means of the procedure laid out in [109]. The values of χ_{exp}^2 and $\chi_{\text{exp+th}}^2$ reported in section 4 are computed accordingly. In this appendix, we recompute the values of $\chi_{\text{exp+th}}^2$ with the original, unregularised experimental covariance matrix. These values, called $\chi_{\text{exp+th,orig}}^2$, are compared to the $\chi_{\text{exp+th}}^2$ values of section 4 in table 9. We also substantiate what the largest changes in covariances and correlations are upon applying our regularisation procedure. To this purpose, we report in table 10, for each of the regularised data sets, the maximum relative difference of the variances $\Delta\sigma_r$ (in percent) and the maximum absolute

Dataset		ABMP16	CT18	CT18A	CT18Z	MSHT20	NNPDF3.1	NNPDF4.0	PDF4LHC15	PDF4LHC21
CMS 13 TeV $W^+ \frac{d\sigma}{d \eta }$	$\chi_{\text{exp+th}}^2$	1.31	1.20	1.11	1.05	1.26	0.85	0.96	1.15	0.98
	$\chi_{\text{exp+th,orig}}^2$	14.6	10.9	10.8	10.8	12.2	10.2	11.2	10.7	10.5
CMS 13 TeV $W^- \frac{d\sigma}{d \eta }$	$\chi_{\text{exp+th}}^2$	1.56	1.15	1.11	1.10	1.43	1.12	1.60	1.14	1.20
	$\chi_{\text{exp+th,orig}}^2$	13.9	8.22	8.32	8.51	10.4	8.26	11.5	8.40	8.59
ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d\sigma}{dm_{t\bar{t}}}$	$\chi_{\text{exp+th}}^2$	0.84	0.99	0.97	0.94	0.97	0.86	0.81	0.96	0.93
	$\chi_{\text{exp+th,orig}}^2$	0.92	1.07	1.05	1.02	1.05	0.94	0.87	1.04	1.01
ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	$\chi_{\text{exp+th}}^2$	0.93	1.38	1.39	1.42	1.48	1.12	1.22	1.22	1.39
	$\chi_{\text{exp+th,orig}}^2$	0.97	1.44	1.44	1.46	1.53	1.17	1.28	1.26	1.45
ATLAS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d\sigma}{dm_{t\bar{t}}}$	$\chi_{\text{exp+th}}^2$	1.41	1.17	1.17	1.04	1.18	1.46	1.39	1.20	1.19
	$\chi_{\text{exp+th,orig}}^2$	1.45	1.24	1.23	1.10	1.24	1.48	1.41	1.26	1.24
ATLAS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d\sigma}{dp_T^t}$	$\chi_{\text{exp+th}}^2$	0.56	0.54	0.54	0.52	0.53	0.56	0.53	0.53	0.53
	$\chi_{\text{exp+th,orig}}^2$	0.80	0.83	0.83	0.81	0.82	0.81	0.78	0.81	0.81
CMS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	$\chi_{\text{exp+th}}^2$	2.77	2.89	2.87	2.76	3.36	3.01	3.61	2.81	2.81
	$\chi_{\text{exp+th,orig}}^2$	3.16	3.29	3.27	3.27	3.80	3.47	4.19	3.21	3.26
ATLAS 13 TeV single-inclusive jets $\frac{d^2\sigma}{dp_T d y }$	$\chi_{\text{exp+th}}^2$	1.85	1.56	1.64	1.38	1.67	1.21	1.51	1.20	1.25
	$\chi_{\text{exp+th,orig}}^2$	3.25	2.77	2.90	2.49	2.86	2.17	2.80	2.28	2.38
CMS 13 TeV single-inclusive jets ($R = 0.4$) $\frac{d^2\sigma}{dp_T d y }$	$\chi_{\text{exp+th}}^2$	1.64	1.58	1.60	1.52	1.64	1.47	1.50	1.48	1.43
	$\chi_{\text{exp+th,orig}}^2$	3.04	2.68	2.69	2.63	2.94	2.48	2.70	2.51	2.62
ATLAS 13 TeV di-jets $\frac{d^2\sigma}{dm_{jj}d y^* }$	$\chi_{\text{exp+th}}^2$	1.13	1.08	1.09	1.05	1.16	1.09	1.15	1.01	0.96
	$\chi_{\text{exp+th,orig}}^2$	1.73	1.56	1.58	1.52	1.72	1.53	1.70	1.46	1.41
H1 single-inclusive-jets (low Q^2) $\frac{d^2\sigma}{dQ^2 dp_T}$	$\chi_{\text{exp+th}}^2$	1.64	1.61	1.61	1.67	1.61	1.70	1.74	1.61	1.73
	$\chi_{\text{exp+th,orig}}^2$	2.20	2.19	2.19	2.27	2.20	2.30	2.34	2.17	2.33
H1 di-jets (low Q^2) $\frac{d^2\sigma}{dQ^2 d(p_T)}$	$\chi_{\text{exp+th}}^2$	1.37	1.39	1.38	1.37	1.39	1.42	1.44	1.36	1.44
	$\chi_{\text{exp+th,orig}}^2$	2.28	2.29	2.29	2.21	2.31	2.32	2.34	2.27	2.34

Table 9. A comparison of the values of $\chi_{\text{exp+th}}^2$, computed in section 4 by regularizing the experimental covariance matrix with the procedure of [109], to the corresponding values $\chi_{\text{exp+th,orig}}^2$, computed with the original, unregularised covariance matrix.

difference of the correlation $|\Delta\rho|$ computed between the corresponding unregularised and regularised matrices.

From table 9, we see that the effect of regularisation on the χ^2 depends on the dataset. For some of these, the effect is huge, e.g. for the CMS Drell-Yan measurement or for the ATLAS and CMS single-inclusive jet and di-jet measurements. Specifically, it amounts to a reduction of the χ^2 of more than 20σ for the first and of about 7σ for the latter. For others, the effect is small, e.g. for the ATLAS and CMS top-quark pair measurements or for the H1 single-inclusive jet measurements. This is unsurprising, given that the first datasets have the largest value of $Z_{\mathcal{L}}$ among all the datasets selected in table 2. One may think

	CMS 13 TeV $W^+ \frac{d\sigma}{d \eta }$	CMS 13 TeV $W^- \frac{d\sigma}{d \eta }$	ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d\sigma}{dm_{t\bar{t}}}$	ATLAS 13 TeV $t\bar{t}$ all hadr. $\frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	ATLAS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d\sigma}{dm_{t\bar{t}}}$	ATLAS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d\sigma}{dp_T^t}$	CMS 13 TeV $t\bar{t} \ell + j \frac{1}{\sigma} \frac{d^2\sigma}{dm_{t\bar{t}}d y_{t\bar{t}} }$	ATLAS 13 TeV single-inclusive jets $\frac{d^2\sigma}{dp_T d y }$	CMS 13 TeV single-inclusive jets ($R = 0.4$) $\frac{d^2\sigma}{dp_T d y }$	ATLAS 13 TeV di-jets $\frac{d^2\sigma}{dm_{jj} d y^* }$	H1 single-inclusive-jets (low Q^2) $\frac{d^2\sigma}{dQ^2 dp_T}$	H1 di-jets (low Q^2) $\frac{d^2\sigma}{dQ^2 d p_T }$
$\Delta\sigma_r$	5.48	5.45	2.54	1.37	3.06	2.71	4.07	5.54	5.41	4.91	2.76	3.73
$ \Delta\rho $	0.06	0.06	0.03	0.02	0.04	0.04	0.06	0.06	0.06	0.06	0.03	0.04

Table 10. The maximum relative difference of the variances $\Delta\sigma_r$ (in percent) and the maximum absolute difference of the correlation $|\Delta\rho|$ computed between the corresponding unregularised and regularised matrices for the regularised datasets outlined in table 9.

that the regularisation procedure is significantly modifying uncertainties and correlations. However, as we can see from table 10, changes are relatively mild: the largest relative change of uncertainties is of order 5%, whereas the largest absolute change of correlation is of order 0.06. The first of this changes results in an effective increase of the diagonal elements of the covariance matrix. The increase is however moderate: typical LHC uncertainties are around a few percent, so the actual increase in uncertainty is of the order of a few permil. The second of this changes results in an effective decrease of correlations by about 6%. We consider this decrease to also be moderate, and at the same time we appreciate that it may be experimentally very challenging to quantify correlations with a precision of 6%.

It may seem counter-intuitive that a relatively small change in the covariance matrix leads to a variation of several standard deviations in the χ^2 . We refer the reader to [109] for a mathematical demonstration of this fact. Here we shall note that the degree of knowledge of experimental correlations related to a 1σ variation of the χ^2 depends on the size of the uncertainties. The smaller the uncertainty, the higher the required degree of precision. Roughly speaking, as one can see from figure 3 of [109], for a 1% (2.5%) uncertainty, correlations ought to be known with a precision of 2% (12%) in order to be within a variation of the χ^2 of one standard deviation. It is therefore unsurprising that the largest improvements in the value of the χ^2 occur for the most precise datasets, which are affected by percent-level (if not sub-percent-level) uncertainties. That being said, we reiterate the fact that the regularisation procedure applies to the covariance matrix as a whole, hence it does not discriminate across different uncertainties that could have a different physical meaning. If we assessed a dataset for inclusion in a PDF determination, understanding which uncertainties are responsible for the bad conditioning of the covariance matrix would be mandatory. However, we consider

that all of these observations do not affect our ability to comparatively judge the performance of different PDF sets at describing the data. In our view, the regularisation procedure does not alter the relative pattern of χ^2 among different PDF sets and datasets. Therefore, the conclusions of section 4 continue to hold.

B Additional results

In this appendix, we collect additional results, complementing those presented in section 4, for the breakdown of $\chi_{\text{exp+th}}^2$ into its $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 components, and for the data-theory comparisons. The additional results refer to the following categories of measurements.

Top-quark pair production. Figure 15 displays the breakdown of $\chi_{\text{exp+th}}^2$ into its $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 components for the datasets not displayed in figure 5, namely: the ATLAS all-hadronic absolute differential distribution in the invariant mass of the top-quark pair; the ATLAS all-hadronic normalised differential distribution in the absolute rapidity of the top-quark pair; the ATLAS lepton+jets normalised differential distributions in the absolute rapidity of the top quark and of the top-quark pair; and the CMS lepton+jets normalised differential distributions in the transverse momentum of the top quark and of the invariant mass of the top-quark pair. Figure 16 displays the data-theory comparison for the top-quark pair single-differential distributions not displayed in figure 6, namely: the ATLAS all-hadronic normalised distribution differential in the absolute value of the top-quark pair rapidity; the ATLAS all-hadronic absolute distribution differential in the invariant mass of the top-quark pair; the ATLAS lepton+jets normalised distributions differential in the absolute rapidity of the top quark and of the top-quark pair; and the CMS lepton+jets normalised distributions differential in the transverse momentum of the top quark and in the invariant mass of the top-quark pair. Figure 3 displays the data-theory comparison for the top-quark pair bins of the ATLAS and CMS double-differential distributions not displayed in figure 6.

Single-inclusive jet and di-jet production at the LHC. Figures 18, 19, and 20 display the data-theory comparison for the remaining rapidity bins not shown in figures 8, 9, and 10, respectively. Figure 18 corresponds to the ATLAS single-inclusive jet measurement; figure 19 corresponds to the CMS single-inclusive jet measurement; and figure 19 corresponds to the ATLAS di-jet measurement.

Single-inclusive jet and di-jet production at HERA. Figure 21 displays the breakdown of $\chi_{\text{exp+th}}^2$ into its $\chi_{\text{exp+mho}}^2$ and χ_{exp}^2 components for the ZEUS single-inclusive jet and di-jet measurements outlined in table 2. Figures 22–25 display the data-theory comparison for the H1 Q^2 bins not displayed in figure 12, respectively, for the low- Q^2 single-inclusive jet and di-jet measurements, and for the high- Q^2 single-inclusive jet and di-jet measurements. Figures 26–28 display the data-theory comparison, respectively, for the ZEUS low-luminosity single-inclusive jet, high-luminosity single-inclusive jet, and di-jet measurements.

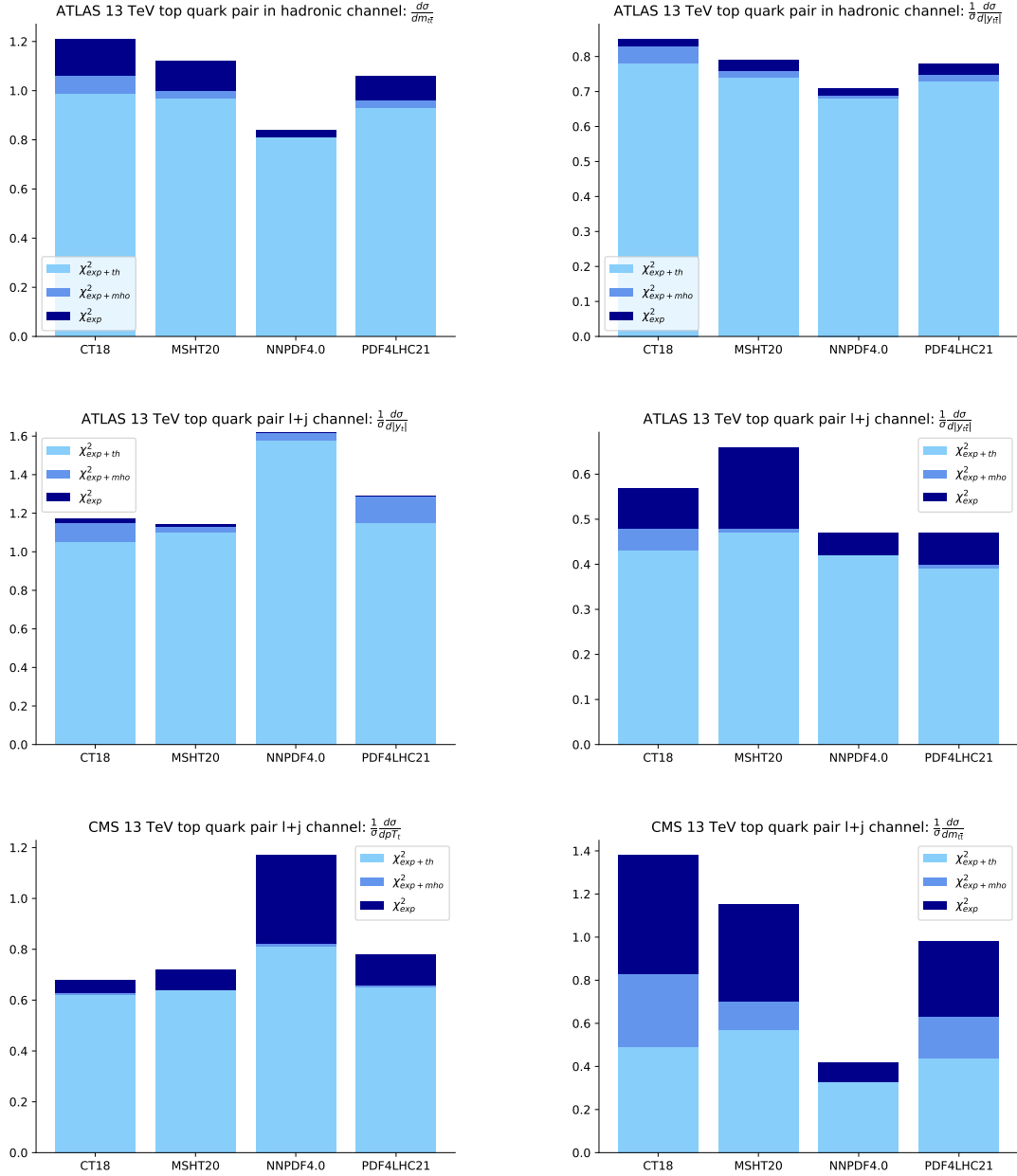


Figure 15. Same as figure 5 for the ATLAS and CMS datasets not shown there. From top to bottom, left to right: the ATLAS all-hadronic absolute differential distribution in the invariant mass of the top-quark pair ($n_{\text{dat}} = 9$, $\sqrt{2/n_{\text{dat}}} = 0.47$); the ATLAS all-hadronic normalised differential distribution in the absolute rapidity of the top-quark pair ($n_{\text{dat}} = 12$, $\sqrt{2/n_{\text{dat}}} = 0.41$); the ATLAS lepton+jets normalised differential distributions in the absolute rapidity of the top quark ($n_{\text{dat}} = 5$, $\sqrt{2/n_{\text{dat}}} = 0.63$) and of the top-quark pair ($n_{\text{dat}} = 7$, $\sqrt{2/n_{\text{dat}}} = 0.53$); and the CMS lepton+jets normalised differential distributions in the transverse momentum of the top quark ($n_{\text{dat}} = 16$, $\sqrt{2/n_{\text{dat}}} = 0.35$) and of the invariant mass of the top-quark pair ($n_{\text{dat}} = 15$, $\sqrt{2/n_{\text{dat}}} = 0.37$).

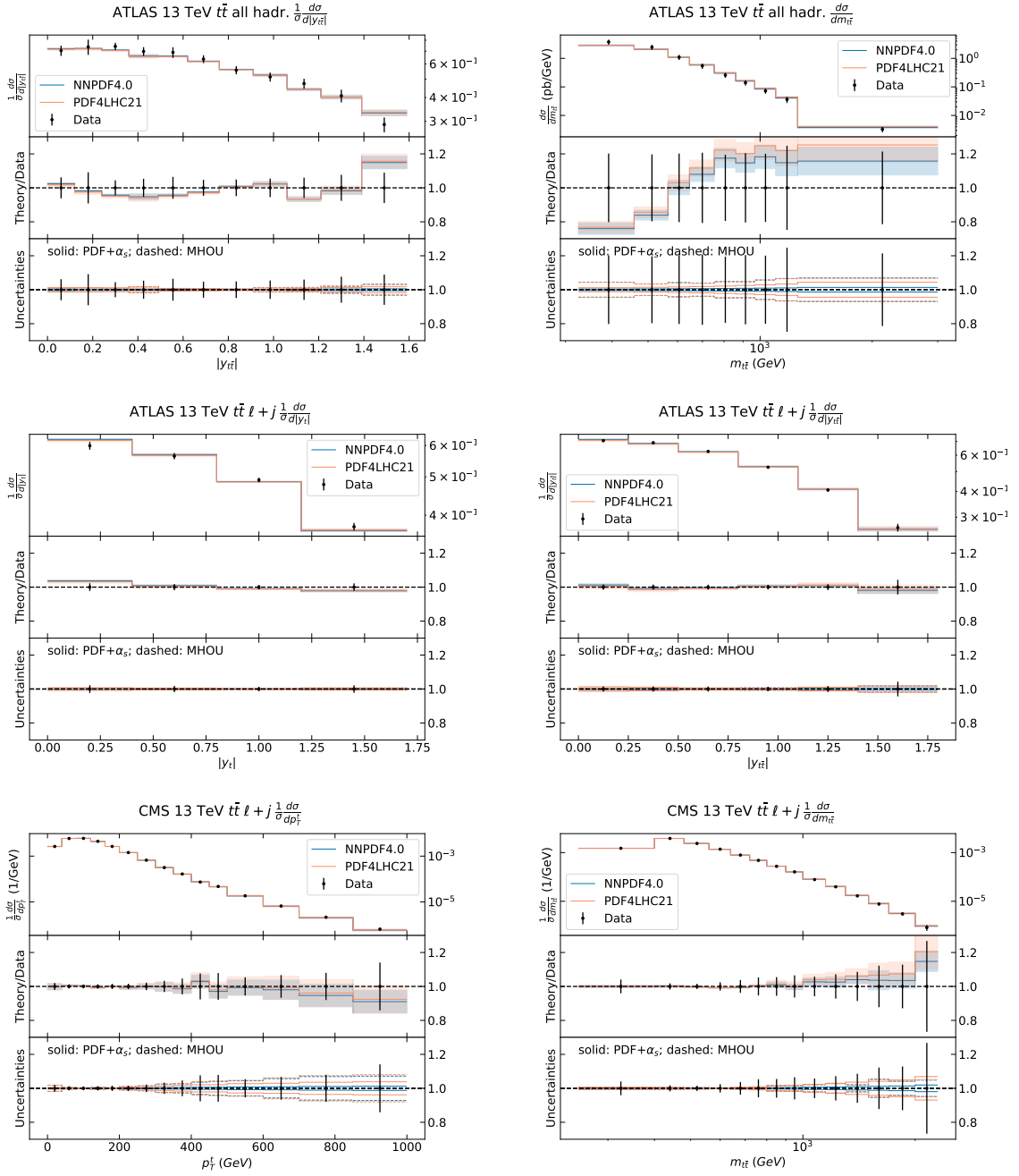


Figure 16. Same as figure 3 for the top-quark pair single-differential distributions not displayed in figure 6, namely: the ATLAS all-hadronic normalised distribution differential in the absolute value of the top-quark pair rapidity; the ATLAS all-hadronic absolute distribution differential in the invariant mass of the top-quark pair; the ATLAS lepton+jets normalised distributions differential in the absolute rapidity of the top quark and of the top-quark pair; and the CMS lepton+jets normalised distributions differential in the transverse momentum of the top quark and in the invariant mass of the top-quark pair.

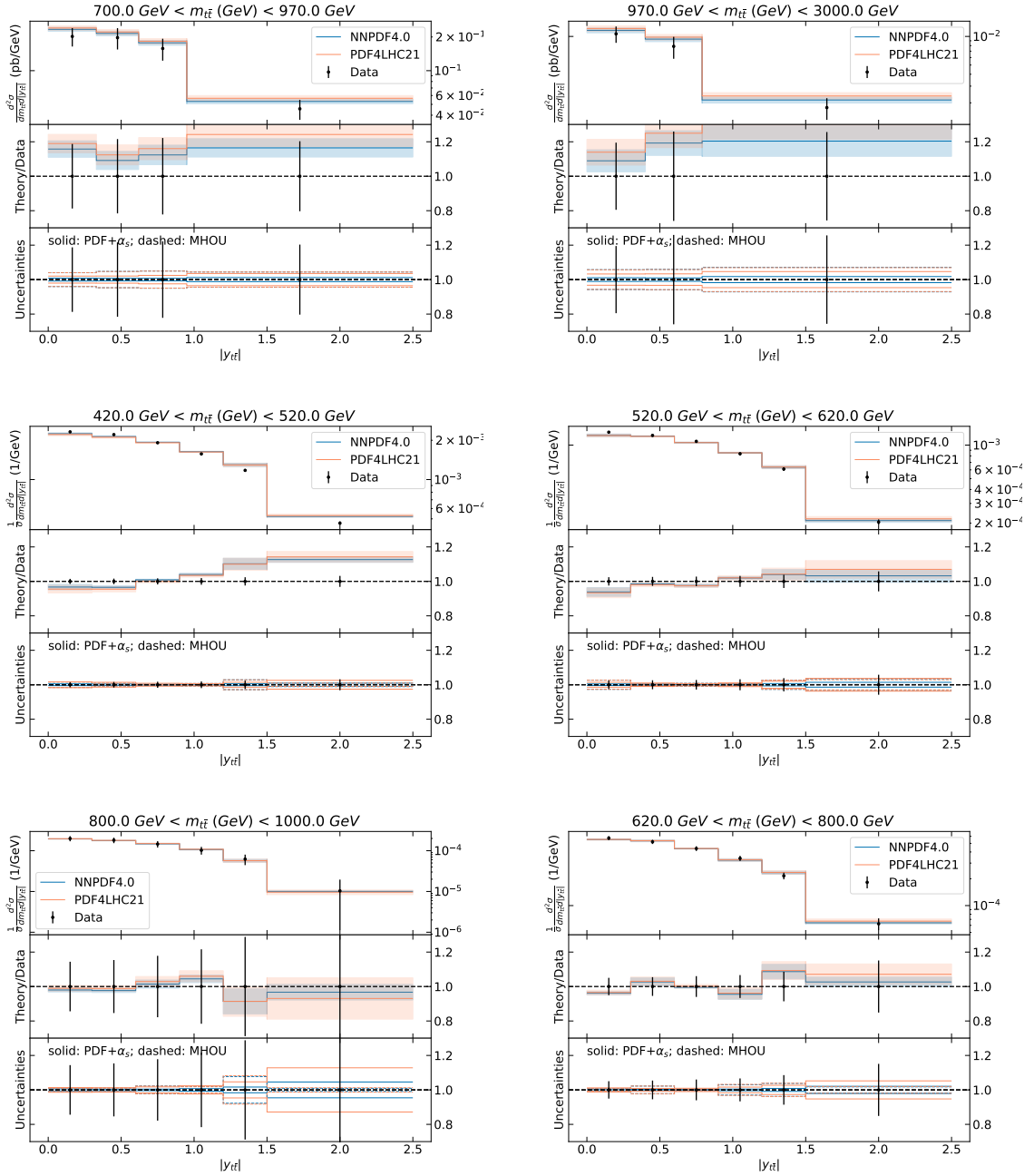


Figure 17. Same as figure 3 for the top-quark pair bins of the double-differential distributions not displayed in figure 6: the first row corresponds to the ATLAS measurement in the all-hadronic final state; the second and third rows correspond to the CMS measurement in the lepton+jets final state.

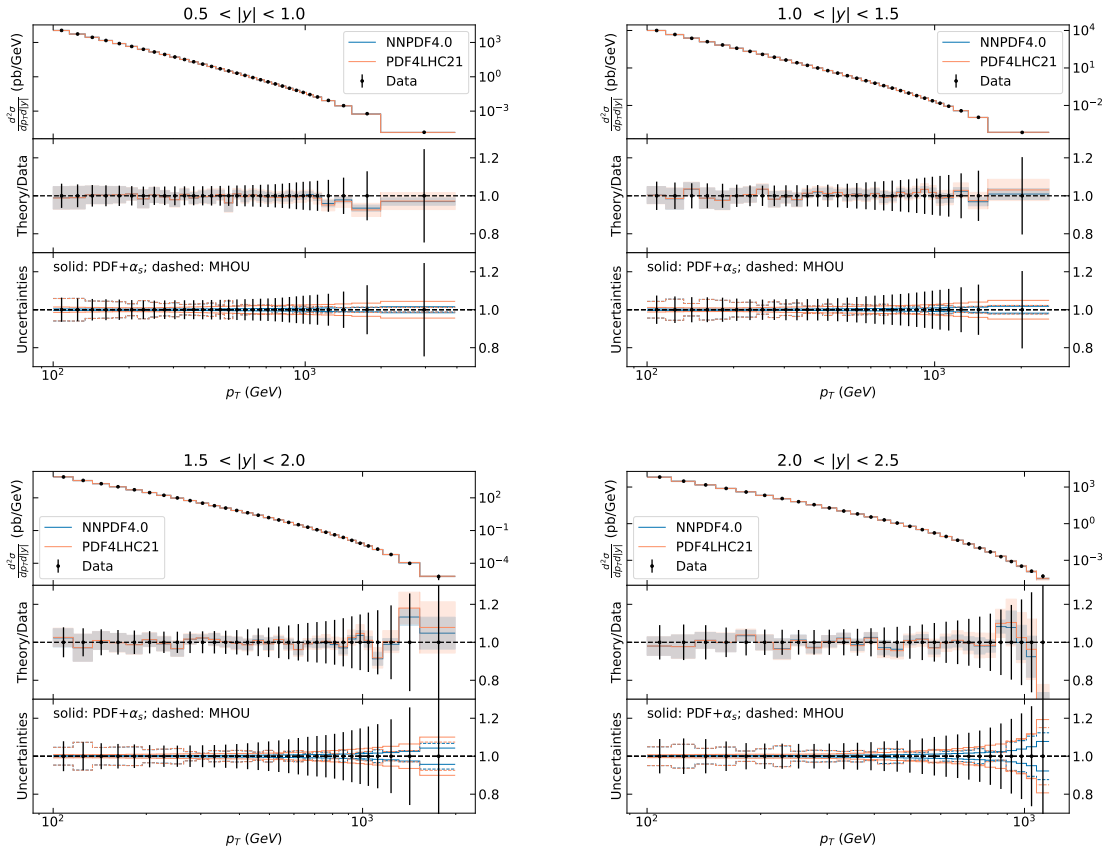


Figure 18. Same as figure 8 for the intermediate rapidity bins.

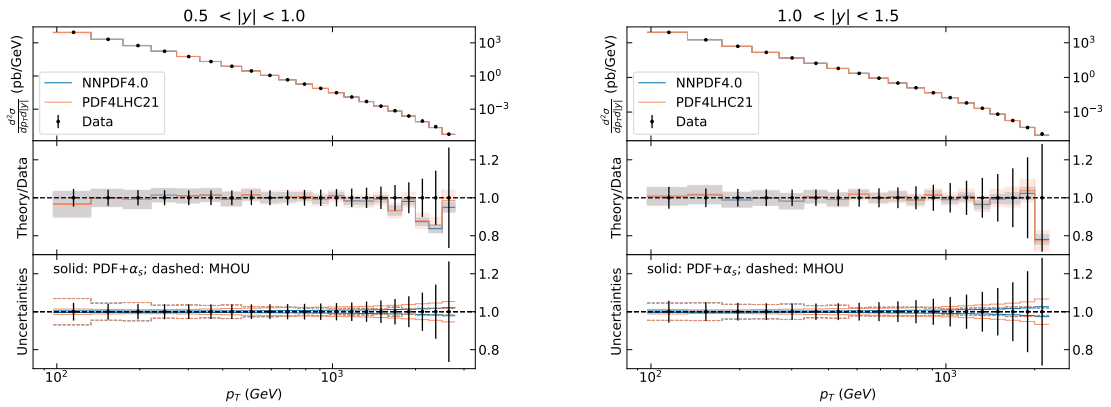


Figure 19. Same as figure 9 for the intermediate rapidity bins.

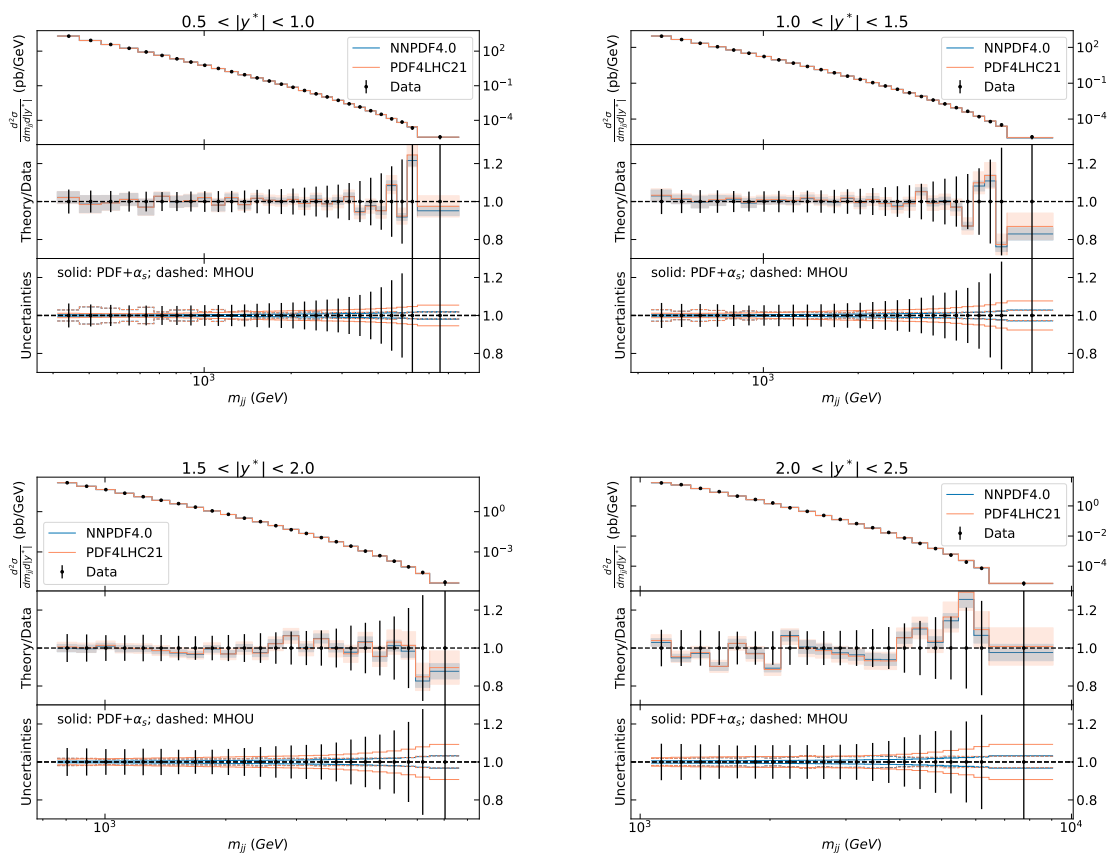


Figure 20. Same as figure 10 for the intermediate rapidity bins.

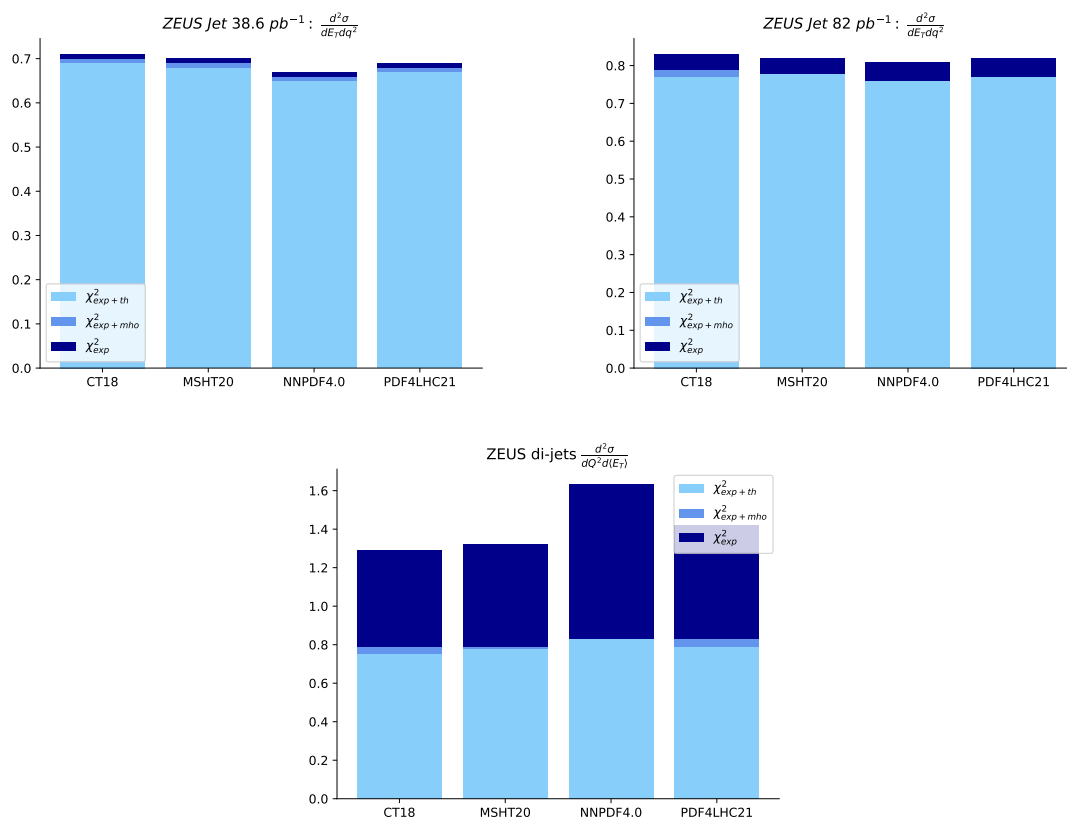


Figure 21. Same as figure 2 for the ZEUS single-inclusive jet dataset comprising $n_{\text{dat}} = 30$ data points (and $\sqrt{2/n_{\text{dat}}} = 0.26$) (top) and di-jet (bottom) dataset comprising $n_{\text{dat}} = 22$ data points (and $\sqrt{2/n_{\text{dat}}} = 0.30$).

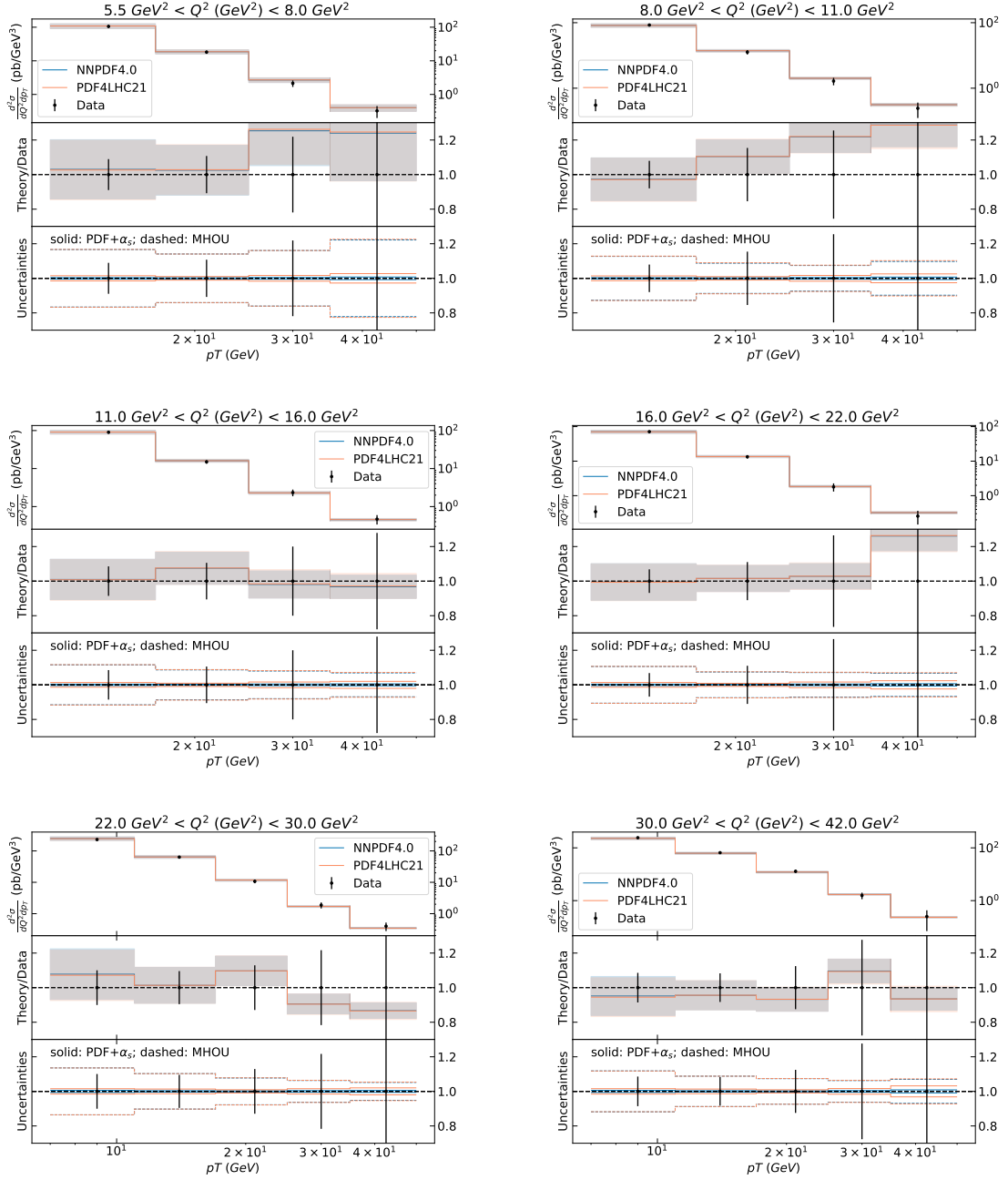


Figure 22. Same as figure 12 for the bins of the H1 low- Q^2 single-inclusive jet measurement not displayed in figure 12.

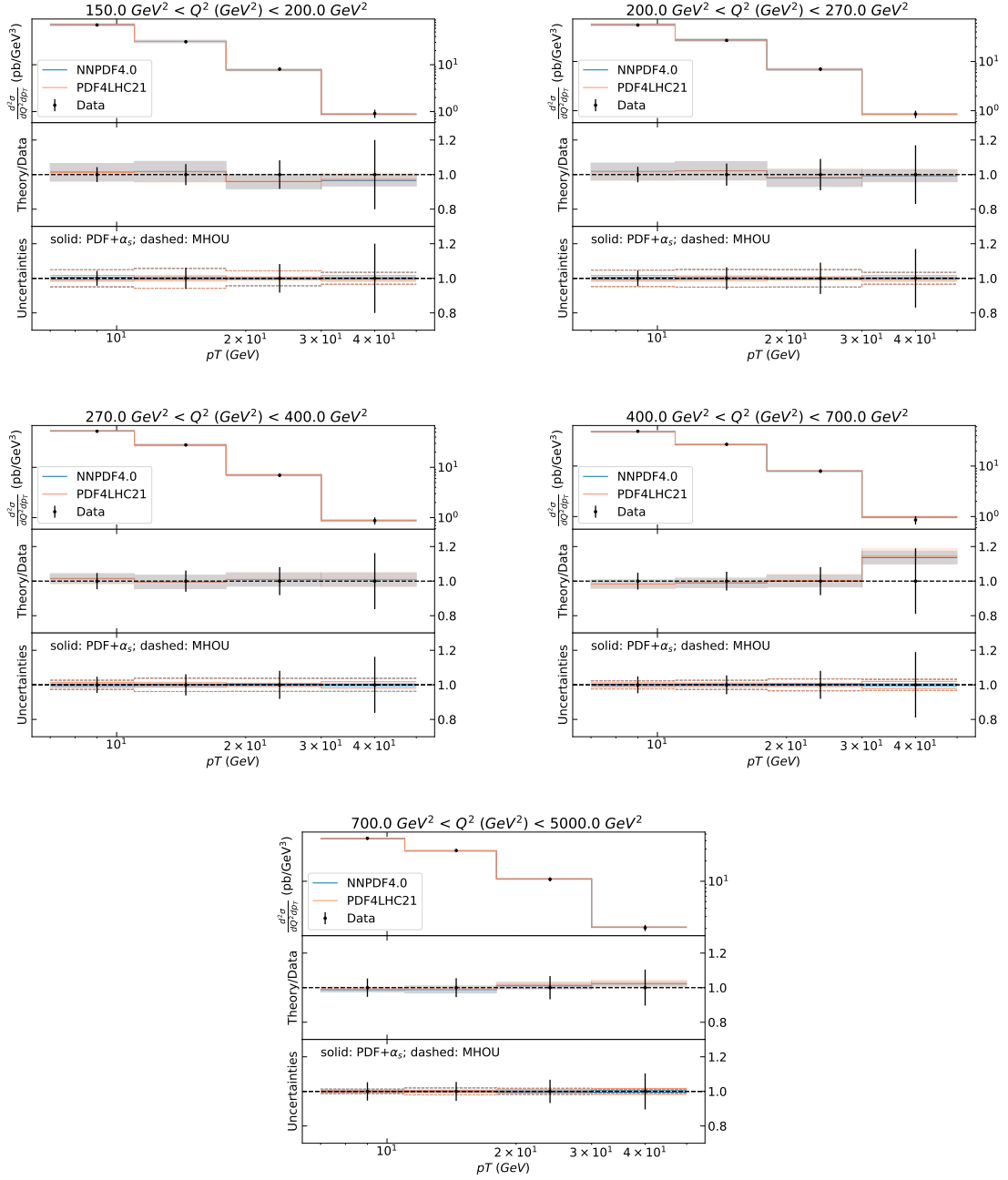


Figure 23. Same as figure 12 for the bins of the H1 high- Q^2 single-inclusive jet measurement not displayed in figure 12.

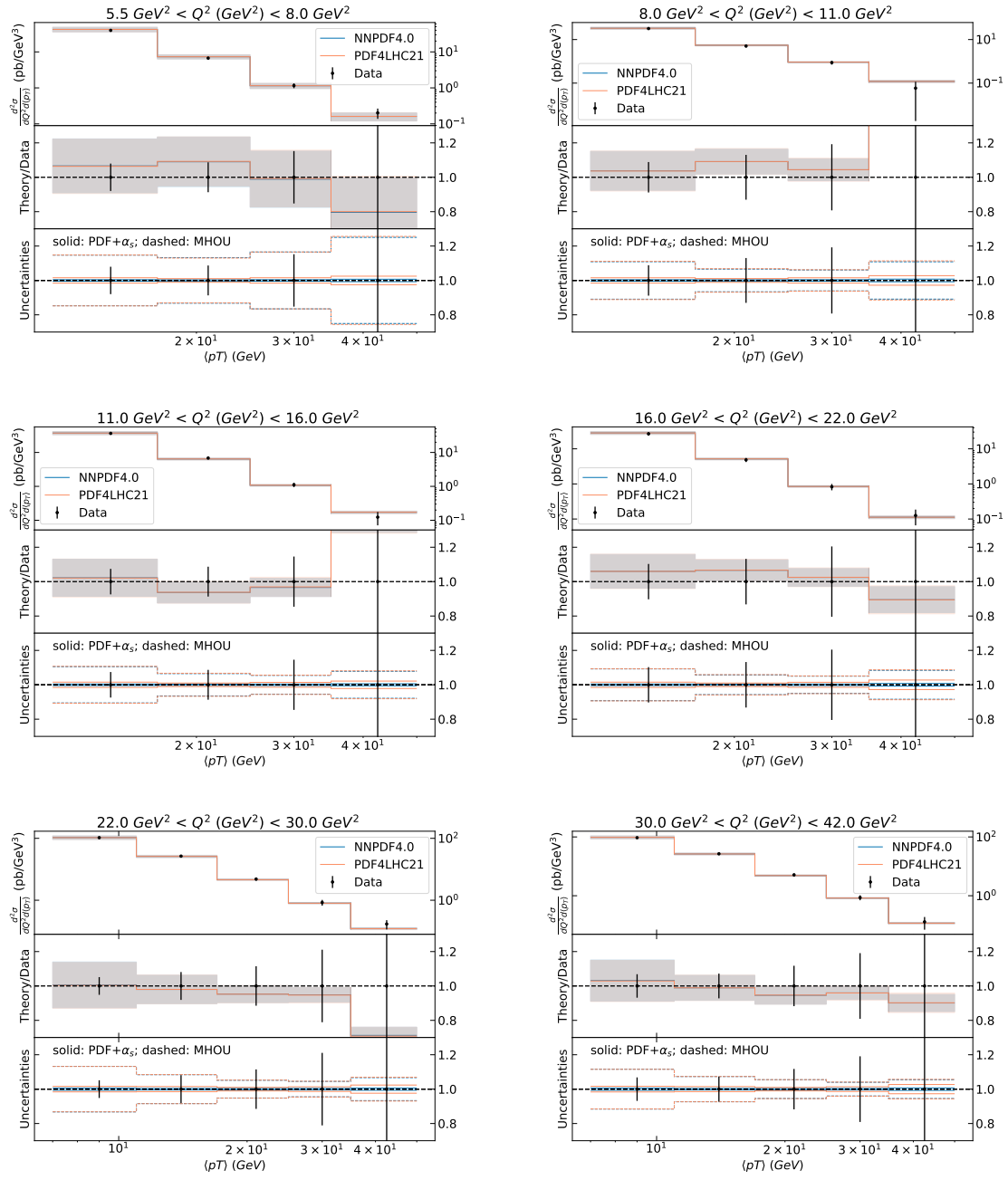


Figure 24. Same as figure 12 for the bins of the H1 low- Q^2 di-jet measurement not displayed in figure 12.

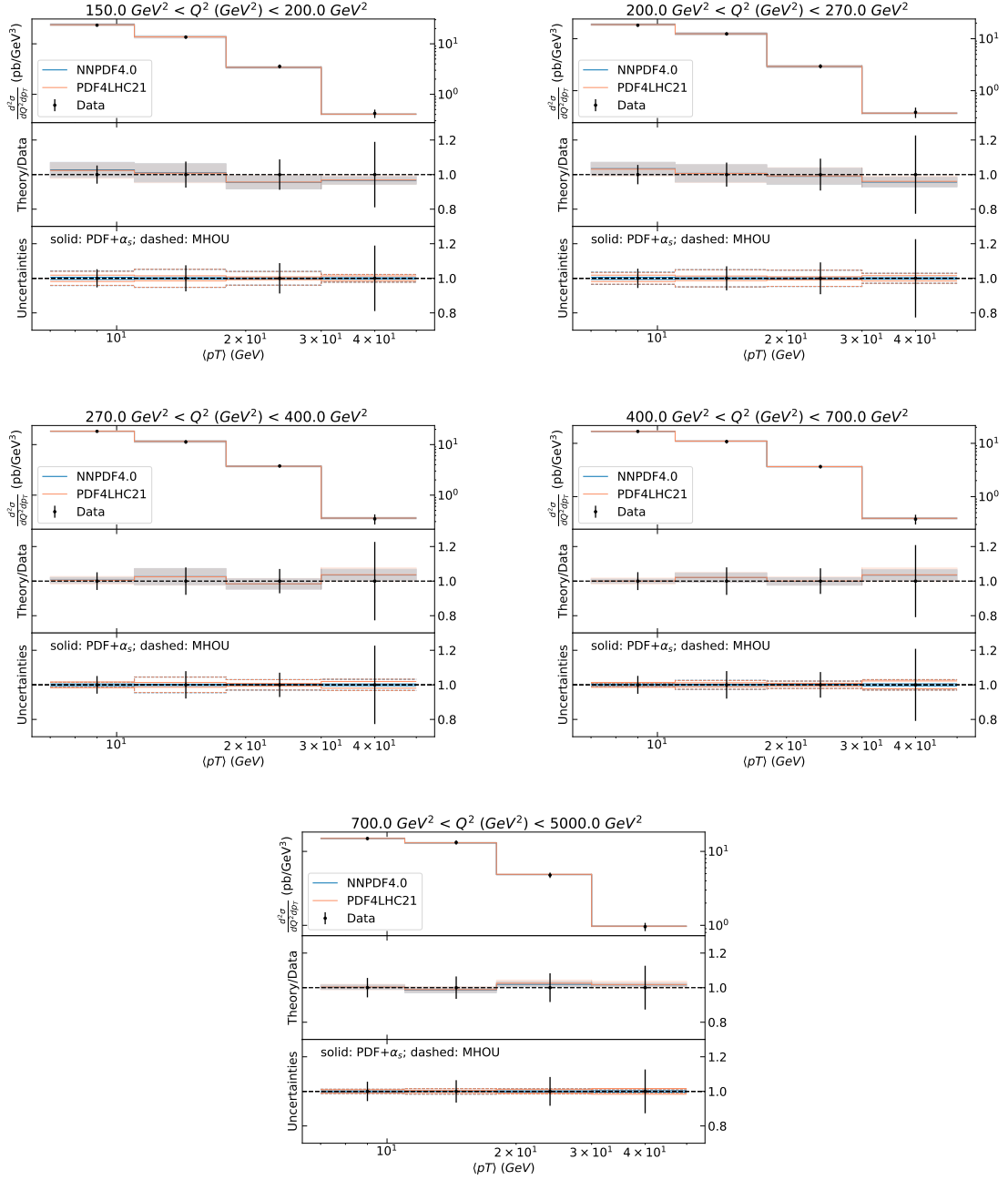


Figure 25. Same as figure 12 for the bins of the H1 high- Q^2 di-jet measurement not displayed in figure 12.

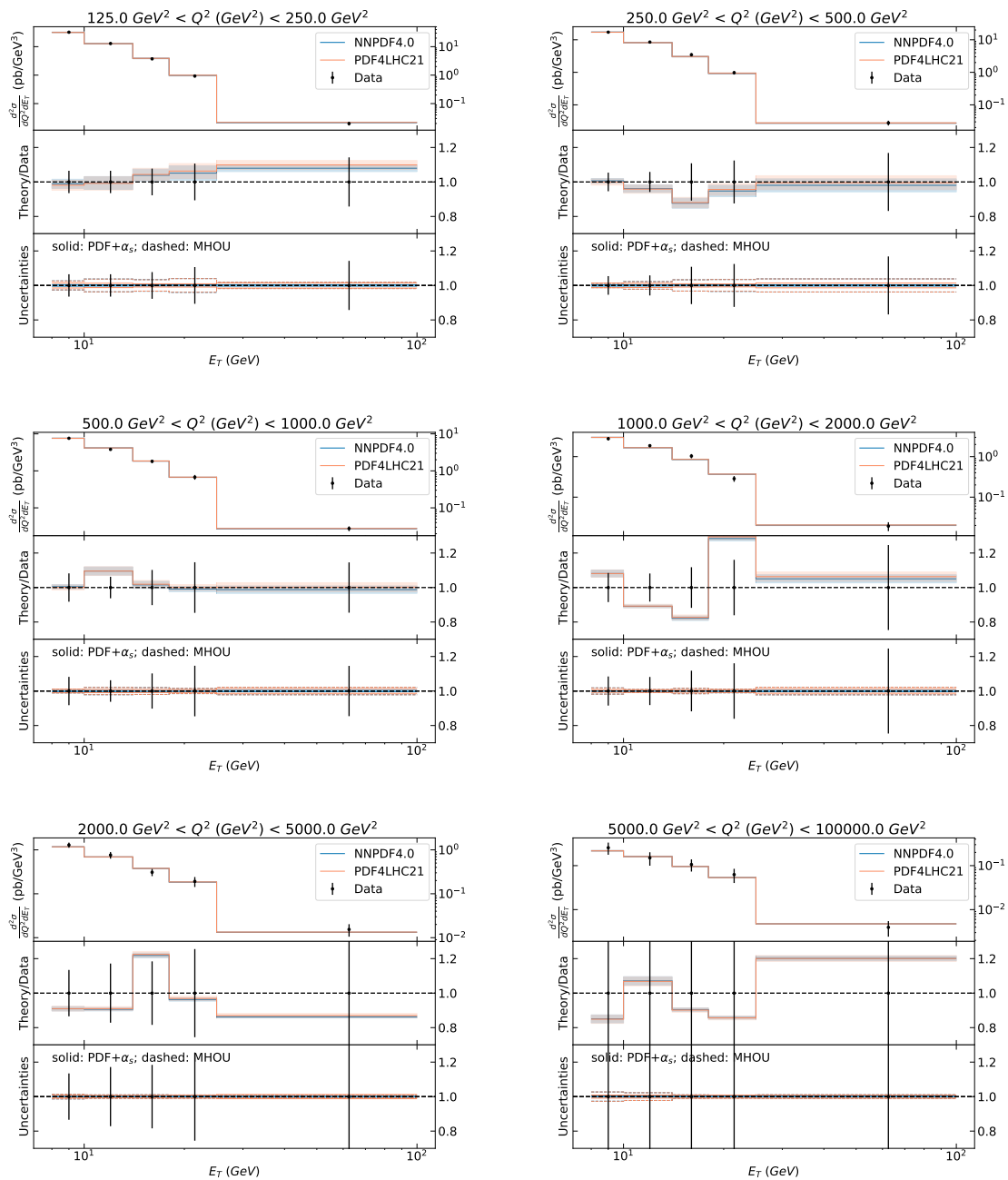


Figure 26. Same as figure 3 for the ZEUS low-luminosity single-inclusive jet production measurement.

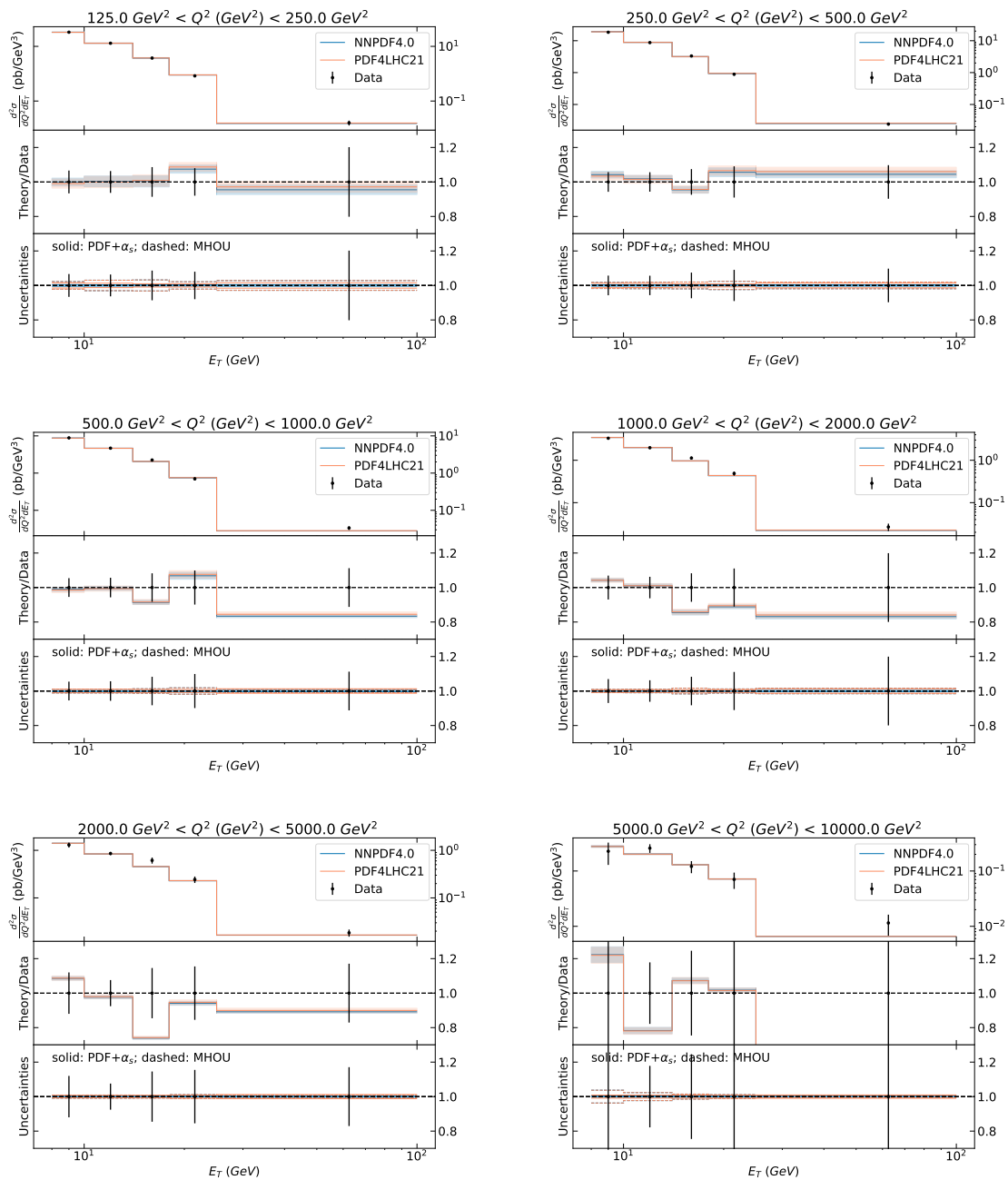


Figure 27. Same as figure 3 for the ZEUS high-luminosity single-inclusive jet production measurement.

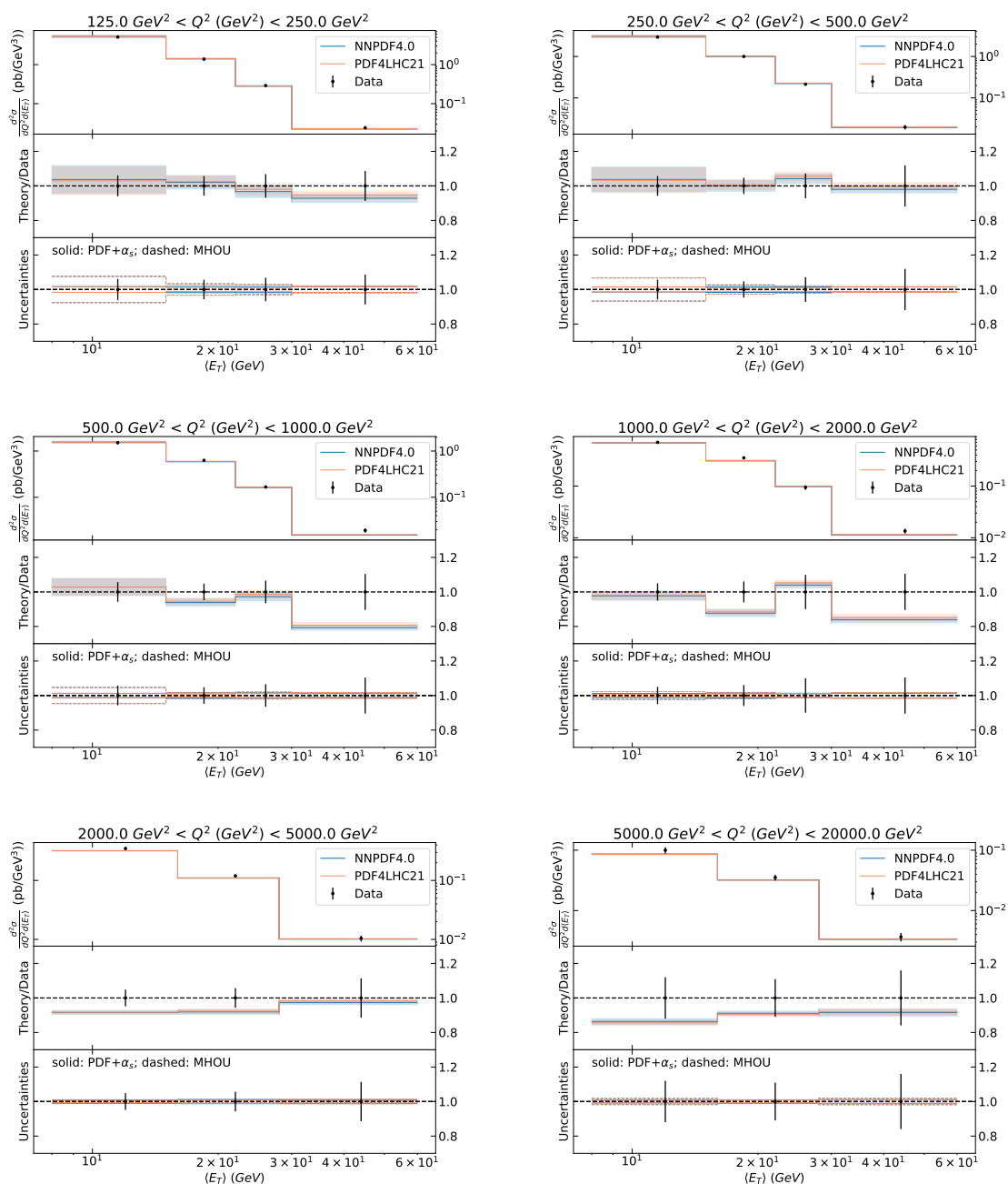


Figure 28. Same as figure 3 for the ZEUS di-jet production measurement.

Data Availability Statement. This article has associated data in a data repository.

Code Availability Statement. This article has associated code in a code repository.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] K. Kovařík, P.M. Nadolsky and D.E. Soper, *Hadronic structure in high-energy collisions*, *Rev. Mod. Phys.* **92** (2020) 045003 [[arXiv:1905.06957](#)] [[INSPIRE](#)].
- [2] J.J. Ethier and E.R. Nocera, *Parton distributions in nucleons and nuclei*, *Ann. Rev. Nucl. Part. Sci.* **70** (2020) 43 [[arXiv:2001.07722](#)] [[INSPIRE](#)].
- [3] S. Amoroso et al., *Snowmass 2021 whitepaper: proton structure at the precision frontier*, *Acta Phys. Polon. B* **53** (2022) 12 [[arXiv:2203.13923](#)] [[INSPIRE](#)].
- [4] M. Ubiali, *Parton distribution functions and their impact on precision of the current theory calculations*, [arXiv:2404.08508](#) [[INSPIRE](#)].
- [5] S. Alekhin, J. Blümlein, S. Moch and R. Placakyte, *Parton distribution functions, α_s , and heavy-quark masses for LHC run II*, *Phys. Rev. D* **96** (2017) 014011 [[arXiv:1701.05838](#)] [[INSPIRE](#)].
- [6] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [[arXiv:1912.10053](#)] [[INSPIRE](#)].
- [7] S. Bailey et al., *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) 341 [[arXiv:2012.04684](#)] [[INSPIRE](#)].
- [8] NNPDF collaboration, *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022) 428 [[arXiv:2109.02653](#)] [[INSPIRE](#)].
- [9] JEFFERSON LAB ANGULAR MOMENTUM (JAM) collaboration, *Simultaneous Monte Carlo analysis of parton densities and fragmentation functions*, *Phys. Rev. D* **104** (2021) 016015 [[arXiv:2101.04664](#)] [[INSPIRE](#)].
- [10] ATLAS collaboration, *Determination of the parton distribution functions of the proton using diverse ATLAS data from pp collisions at $\sqrt{s} = 7, 8$ and 13 TeV*, *Eur. Phys. J. C* **82** (2022) 438 [[arXiv:2112.11266](#)] [[INSPIRE](#)].
- [11] A. Accardi, X. Jing, J.F. Owens and S. Park, *Light quark and antiquark constraints from new electroweak data*, *Phys. Rev. D* **107** (2023) 113005 [[arXiv:2303.11509](#)] [[INSPIRE](#)].
- [12] S. Alekhin et al., *The PDF4LHC working group interim report*, [arXiv:1101.0536](#) [[INSPIRE](#)].
- [13] M. Botje et al., *The PDF4LHC working group interim recommendations*, [arXiv:1101.0538](#) [[INSPIRE](#)].
- [14] R.D. Ball et al., *Parton distribution benchmarking with LHC data*, *JHEP* **04** (2013) 125 [[arXiv:1211.5142](#)] [[INSPIRE](#)].
- [15] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: results from run I and preparation for run II*, *J. Phys. G* **42** (2015) 103103 [[arXiv:1507.00556](#)] [[INSPIRE](#)].
- [16] A. Accardi et al., *A critical appraisal and evaluation of modern PDFs*, *Eur. Phys. J. C* **76** (2016) 471 [[arXiv:1603.08906](#)] [[INSPIRE](#)].
- [17] J.R. Andersen et al., *Les Houches 2015: physics at TeV colliders Standard Model working group report*, in the proceedings of the 9th Les Houches workshop on physics at TeV colliders, (2016) [[arXiv:1605.04692](#)] [[INSPIRE](#)].
- [18] J.R. Andersen et al., *Les Houches 2013: physics at TeV colliders. Standard Model working group report*, [arXiv:1405.1067](#) [[INSPIRE](#)].
- [19] J. Butterworth et al., *PDF4LHC recommendations for LHC run II*, *J. Phys. G* **43** (2016) 023001 [[arXiv:1510.03865](#)] [[INSPIRE](#)].

- [20] G. Watt and R.S. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs*, *JHEP* **08** (2012) 052 [[arXiv:1205.4024](#)] [[INSPIRE](#)].
- [21] T.-J. Hou et al., *Reconstruction of Monte Carlo replicas from Hessian parton distributions*, *JHEP* **03** (2017) 099 [[arXiv:1607.06066](#)] [[INSPIRE](#)].
- [22] J. Gao and P. Nadolsky, *A meta-analysis of parton distribution functions*, *JHEP* **07** (2014) 035 [[arXiv:1401.0013](#)] [[INSPIRE](#)].
- [23] S. Carrazza et al., *An unbiased Hessian representation for Monte Carlo PDFs*, *Eur. Phys. J. C* **75** (2015) 369 [[arXiv:1505.06736](#)] [[INSPIRE](#)].
- [24] S. Carrazza, S. Forte, Z. Kassabov and J. Rojo, *Specialized minimal PDFs for optimized LHC calculations*, *Eur. Phys. J. C* **76** (2016) 205 [[arXiv:1602.00005](#)] [[INSPIRE](#)].
- [25] S. Carrazza, J.I. Latorre, J. Rojo and G. Watt, *A compression algorithm for the combination of PDF sets*, *Eur. Phys. J. C* **75** (2015) 474 [[arXiv:1504.06469](#)] [[INSPIRE](#)].
- [26] PDF4LHC WORKING GROUP collaboration, *The PDF4LHC21 combination of global PDF fits for the LHC run III*, *J. Phys. G* **49** (2022) 080501 [[arXiv:2203.05506](#)] [[INSPIRE](#)].
- [27] X. Jing et al., *Quantifying the interplay of experimental constraints in analyses of parton distributions*, *Phys. Rev. D* **108** (2023) 034029 [[arXiv:2306.03918](#)] [[INSPIRE](#)].
- [28] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [[arXiv:1706.00428](#)] [[INSPIRE](#)].
- [29] NNPDF collaboration, *Evidence for intrinsic charm quarks in the proton*, *Nature* **608** (2022) 483 [[arXiv:2208.08372](#)] [[INSPIRE](#)].
- [30] NNPDF collaboration, *Intrinsic charm quark valence distribution of the proton*, *Phys. Rev. D* **109** (2024) L091501 [[arXiv:2311.00743](#)] [[INSPIRE](#)].
- [31] M. Guzzi et al., *The persistent nonperturbative charm enigma*, *Phys. Lett. B* **843** (2023) 137975 [[arXiv:2211.01387](#)] [[INSPIRE](#)].
- [32] R.D. Ball et al., *Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study*, *Eur. Phys. J. C* **82** (2022) 1160 [[arXiv:2209.08115](#)] [[INSPIRE](#)].
- [33] Y. Fu, R. Brock, D. Hayden and C.-P. Yuan, *Probing parton distribution functions at large x via Drell–Yan forward–backward asymmetry*, *Phys. Rev. D* **109** (2024) 054006 [[arXiv:2307.07839](#)] [[INSPIRE](#)].
- [34] J. Fiaschi et al., *Z' -boson dilepton searches and the high- x quark density*, *Phys. Lett. B* **841** (2023) 137915 [[arXiv:2211.06188](#)] [[INSPIRE](#)].
- [35] A. Greljo et al., *Parton distributions in the SMEFT from high-energy Drell–Yan tails*, *JHEP* **07** (2021) 122 [[arXiv:2104.02723](#)] [[INSPIRE](#)].
- [36] S. Iranipour and M. Ubiali, *A new generation of simultaneous fits to LHC data using deep learning*, *JHEP* **05** (2022) 032 [[arXiv:2201.07240](#)] [[INSPIRE](#)].
- [37] Z. Kassabov et al., *The top quark legacy of the LHC run II for PDF and SMEFT analyses*, *JHEP* **05** (2023) 205 [[arXiv:2303.06159](#)] [[INSPIRE](#)].
- [38] E. Hammou et al., *Hide and seek: how PDFs can conceal new physics*, *JHEP* **11** (2023) 090 [[arXiv:2307.10370](#)] [[INSPIRE](#)].
- [39] PBSP collaboration, *SIMUnet: an open-source tool for simultaneous global fits of EFT Wilson coefficients and PDFs*, *Eur. Phys. J. C* **84** (2024) 805 [[arXiv:2402.03308](#)] [[INSPIRE](#)].

- [40] E. Hammou and M. Ubiali, *Unravelling new physics signals at the HL-LHC with EIC and FPF constraints*, *Phys. Rev. D* **111** (2025) 095028 [[arXiv:2410.00963](#)] [[INSPIRE](#)].
- [41] J. Gao et al., *Simultaneous CTEQ-TEA extraction of PDFs and SMEFT parameters from jet and $t\bar{t}$ data*, *JHEP* **05** (2023) 003 [[arXiv:2211.01094](#)] [[INSPIRE](#)].
- [42] ATLAS collaboration, *A precise measurement of the Z-boson double-differential transverse momentum and rapidity distributions in the full phase space of the decay leptons with the ATLAS experiment at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J. C* **84** (2024) 315 [[arXiv:2309.09318](#)] [[INSPIRE](#)].
- [43] J. McGowan, T. Cridge, L.A. Harland-Lang and R.S. Thorne, *Approximate N^3 LO parton distribution functions with theoretical uncertainties: MSHT20a N^3 LO PDFs*, *Eur. Phys. J. C* **83** (2023) 185 [Erratum *ibid.* **83** (2023) 302] [[arXiv:2207.04739](#)] [[INSPIRE](#)].
- [44] CMS collaboration, *Measurement of the Drell–Yan forward-backward asymmetry and of the effective leptonic weak mixing angle in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Lett. B* **866** (2025) 139526 [[arXiv:2408.07622](#)] [[INSPIRE](#)].
- [45] ATLAS collaboration, *Improved W boson mass measurement using 7 TeV proton-proton collisions with the ATLAS detector*, [ATLAS-CONF-2023-004](#), CERN, Geneva, Switzerland (2023) [[INSPIRE](#)].
- [46] CMS collaboration, *Measurement of the W boson mass in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [CMS-PAS-SMP-23-002](#), CERN, Geneva, Switzerland (2024) [[INSPIRE](#)].
- [47] S. Forte and Z. Kassabov, *Why α_s cannot be determined from hadronic processes without simultaneously determining the parton distributions*, *Eur. Phys. J. C* **80** (2020) 182 [[arXiv:2001.04986](#)] [[INSPIRE](#)].
- [48] J. Cruz-Martinez, S. Forte and E.R. Nocera, *Future tests of parton distributions*, *Acta Phys. Polon. B* **52** (2021) 243 [[arXiv:2103.08606](#)] [[INSPIRE](#)].
- [49] E. Maguire, L. Heinrich and G. Watt, *HEPData: a repository for high energy physics data*, *J. Phys. Conf. Ser.* **898** (2017) 102006 [[arXiv:1704.05473](#)] [[INSPIRE](#)].
- [50] ATLAS collaboration, *Measurement of the transverse momentum distribution of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 616 [[arXiv:1912.02844](#)] [[INSPIRE](#)].
- [51] CMS collaboration, *Measurements of the W boson rapidity, helicity, double-differential cross sections, and charge asymmetry in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **102** (2020) 092012 [[arXiv:2008.04174](#)] [[INSPIRE](#)].
- [52] LHCb collaboration, *Precision measurement of forward Z boson production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **07** (2022) 026 [[arXiv:2112.07458](#)] [[INSPIRE](#)].
- [53] ATLAS collaboration, *Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *JHEP* **01** (2021) 033 [[arXiv:2006.09274](#)] [[INSPIRE](#)].
- [54] ATLAS collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ +jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 1028 [Erratum *ibid.* **80** (2020) 1092] [[arXiv:1908.07305](#)] [[INSPIRE](#)].
- [55] CMS collaboration, *Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **104** (2021) 092013 [[arXiv:2108.02803](#)] [[INSPIRE](#)].

- [56] ATLAS collaboration, *Measurement of inclusive jet and dijet cross-sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **05** (2018) 195 [[arXiv:1711.02692](#)] [[INSPIRE](#)].
- [57] CMS collaboration, *Measurement and QCD analysis of double-differential inclusive jet cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **02** (2022) 142 [*Addendum ibid.* **12** (2022) 035] [[arXiv:2111.10431](#)] [[INSPIRE](#)].
- [58] H1 collaboration, *Measurement of jet production cross sections in deep-inelastic ep scattering at HERA*, *Eur. Phys. J. C* **77** (2017) 215 [*Erratum ibid.* **81** (2021) 739] [[arXiv:1611.03421](#)] [[INSPIRE](#)].
- [59] H1 collaboration, *Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s* , *Eur. Phys. J. C* **75** (2015) 65 [[arXiv:1406.4709](#)] [[INSPIRE](#)].
- [60] ZEUS collaboration, *Inclusive jet cross-sections in the Breit frame in neutral current deep inelastic scattering at HERA and determination of α_s* , *Phys. Lett. B* **547** (2002) 164 [[hep-ex/0208037](#)] [[INSPIRE](#)].
- [61] ZEUS collaboration, *Inclusive-jet and dijet cross-sections in deep inelastic scattering at HERA*, *Nucl. Phys. B* **765** (2007) 1 [[hep-ex/0608048](#)] [[INSPIRE](#)].
- [62] ZEUS collaboration, *Inclusive dijet cross sections in neutral current deep inelastic scattering at HERA*, *Eur. Phys. J. C* **70** (2010) 965 [[arXiv:1010.6167](#)] [[INSPIRE](#)].
- [63] NNPDF collaboration, *Improving quark flavor separation with forward W and Z production at LHCb*, *PoS DIS2017* (2018) 198 [[arXiv:1705.04468](#)] [[INSPIRE](#)].
- [64] R.S. Thorne, A.D. Martin, W.J. Stirling and G. Watt, *Parton distributions and QCD at LHCb*, in the proceedings of the *16th International Workshop on Deep Inelastic Scattering and Related Subjects*, (2008) [[DOI:10.3360/dis.2008.30](#)] [[arXiv:0808.1847](#)] [[INSPIRE](#)].
- [65] R. Boughezal, A. Guffanti, F. Petriello and M. Ubiali, *The impact of the LHC Z-boson transverse momentum data on PDF determinations*, *JHEP* **07** (2017) 130 [[arXiv:1705.00343](#)] [[INSPIRE](#)].
- [66] LHCb collaboration, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **09** (2016) 136 [[arXiv:1607.06495](#)] [[INSPIRE](#)].
- [67] ATLAS collaboration, *Measurement of the double-differential high-mass Drell-Yan cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **08** (2016) 009 [[arXiv:1606.01736](#)] [[INSPIRE](#)].
- [68] ATLAS collaboration, *Measurement of the Drell-Yan triple-differential cross section in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **12** (2017) 059 [[arXiv:1710.05167](#)] [[INSPIRE](#)].
- [69] S. Carrazza, E.R. Nocera, C. Schwan and M. Zaro, *PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes*, *JHEP* **12** (2020) 108 [[arXiv:2008.12789](#)] [[INSPIRE](#)].
- [70] NNLOJET collaboration, *NNLOJET: a parton-level event generator for jet cross sections at NNLO QCD accuracy*, [arXiv:2503.22804](#) [[INSPIRE](#)].
- [71] J. Cruz-Martinez, A. Huss and C. Schwan, *Fast interpolation grids for the Drell-Yan process*, *Eur. Phys. J. C* **85** (2025) 459 [[arXiv:2501.13167](#)] [[INSPIRE](#)].

- [72] A. Gehrmann-De Ridder et al., *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, *Phys. Rev. Lett.* **117** (2016) 022001 [[arXiv:1507.02850](#)] [[INSPIRE](#)].
- [73] A. Gehrmann-De Ridder et al., *The NNLO QCD corrections to Z boson production at large transverse momentum*, *JHEP* **07** (2016) 133 [[arXiv:1605.04295](#)] [[INSPIRE](#)].
- [74] M. Wiesemann, L. Rottoli and P. Torrielli, *The $Z\gamma$ transverse-momentum spectrum at NNLO+N³LL*, *Phys. Lett. B* **809** (2020) 135718 [[arXiv:2006.09338](#)] [[INSPIRE](#)].
- [75] M. Czakon, M.L. Mangano, A. Mitov and J. Rojo, *Constraints on the gluon PDF from top quark pair production at hadron colliders*, *JHEP* **07** (2013) 167 [[arXiv:1303.7215](#)] [[INSPIRE](#)].
- [76] M. Czakon et al., *Pinning down the large-x gluon with NNLO top-quark pair differential distributions*, *JHEP* **04** (2017) 044 [[arXiv:1611.08609](#)] [[INSPIRE](#)].
- [77] S. Bailey and L. Harland-Lang, *Differential top quark pair production at the LHC: challenges for PDF fits*, *Eur. Phys. J. C* **80** (2020) 60 [[arXiv:1909.10541](#)] [[INSPIRE](#)].
- [78] S. Catani et al., *Top-quark pair hadroproduction at next-to-next-to-leading order in QCD*, *Phys. Rev. D* **99** (2019) 051501 [[arXiv:1901.04005](#)] [[INSPIRE](#)].
- [79] S. Catani et al., *Top-quark pair production at the LHC: fully differential QCD predictions at NNLO*, *JHEP* **07** (2019) 100 [[arXiv:1906.06535](#)] [[INSPIRE](#)].
- [80] S. Catani et al., *Top-quark pair hadroproduction at NNLO: differential predictions with the \overline{MS} mass*, *JHEP* **08** (2020) 027 [[arXiv:2005.00557](#)] [[INSPIRE](#)].
- [81] S. Alekhin, M.V. Garzelli, S.-O. Moch and O. Zenaiev, *NNLO PDFs driven by top-quark data*, *Eur. Phys. J. C* **85** (2025) 162 [[arXiv:2407.00545](#)] [[INSPIRE](#)].
- [82] CMS collaboration, *Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Lett. B* **728** (2014) 496 [*Erratum ibid.* **738** (2014) 526] [[arXiv:1307.1907](#)] [[INSPIRE](#)].
- [83] ATLAS and CMS collaborations, *Combination of inclusive top-quark pair production cross-section measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **07** (2023) 213 [[arXiv:2205.13830](#)] [[INSPIRE](#)].
- [84] M. Czakon, A. Mitov and R. Poncelet, *NNLO QCD corrections to leptonic observables in top-quark pair production and decay*, *JHEP* **05** (2021) 212 [[arXiv:2008.11133](#)] [[INSPIRE](#)].
- [85] M. Grazzini, S. Kallweit and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, *Eur. Phys. J. C* **78** (2018) 537 [[arXiv:1711.06631](#)] [[INSPIRE](#)].
- [86] M. Czakon, D. Heymes and A. Mitov, *Dynamical scales for multi-TeV top-pair production at the LHC*, *JHEP* **04** (2017) 071 [[arXiv:1606.03350](#)] [[INSPIRE](#)].
- [87] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001 [[INSPIRE](#)].
- [88] M. Czakon, D. Heymes and A. Mitov, *fastNLO tables for NNLO top-quark pair differential distributions*, [arXiv:1704.08551](#) [[INSPIRE](#)].
- [89] M. Czakon, D. Heymes and A. Mitov, *High-precision differential predictions for top-quark pairs at the LHC*, *Phys. Rev. Lett.* **116** (2016) 082003 [[arXiv:1511.00549](#)] [[INSPIRE](#)].
- [90] J. Rojo, *Constraints on parton distributions and the strong coupling from LHC jet data*, *Int. J. Mod. Phys. A* **30** (2015) 1546005 [[arXiv:1410.7728](#)] [[INSPIRE](#)].

- [91] R. Abdul Khalek et al., *Phenomenology of NNLO jet production at the LHC and its impact on parton distributions*, *Eur. Phys. J. C* **80** (2020) 797 [[arXiv:2005.11327](#)] [[INSPIRE](#)].
- [92] T. Cridge, L.A. Harland-Lang and R.S. Thorne, *The impact of LHC jet and Z p_T data at up to approximate N^3LO order in the MSHT global PDF fit*, *Eur. Phys. J. C* **84** (2024) 446 [[arXiv:2312.12505](#)] [[INSPIRE](#)].
- [93] CTEQ-TEA collaboration, *Impact of LHC precision measurements of inclusive jet and dijet production on the CTEQ-TEA global PDF fit*, *Phys. Rev. D* **111** (2025) 036033 [[arXiv:2412.00350](#)] [[INSPIRE](#)].
- [94] M. Cacciari, G.P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[INSPIRE](#)].
- [95] A. Gehrmann-De Ridder et al., *Triple differential dijet cross section at the LHC*, *Phys. Rev. Lett.* **123** (2019) 102001 [[arXiv:1905.09047](#)] [[INSPIRE](#)].
- [96] D. Britzger et al., *NNLO interpolation grids for jet production at the LHC*, *Eur. Phys. J. C* **82** (2022) 930 [[arXiv:2207.13735](#)] [[INSPIRE](#)].
- [97] *Ploughshare webpage*, <https://ploughshare.web.cern.ch/ploughshare/>.
- [98] ZEUS collaboration, *Measurement of jet production in deep inelastic scattering and NNLO determination of the strong coupling at ZEUS*, *Eur. Phys. J. C* **83** (2023) 1082 [[arXiv:2309.02889](#)] [[INSPIRE](#)].
- [99] H1 and ZEUS collaborations, *Impact of jet-production data on the next-to-next-to-leading-order determination of HERAPDF2.0 parton distributions*, *Eur. Phys. J. C* **82** (2022) 243 [[arXiv:2112.01120](#)] [[INSPIRE](#)].
- [100] H1 collaboration, *Determination of the strong coupling constant $\alpha_s(m_Z)$ in next-to-next-to-leading order QCD using H1 jet cross section measurements*, *Eur. Phys. J. C* **77** (2017) 791 [*Erratum ibid.* **81** (2021) 738] [[arXiv:1709.07251](#)] [[INSPIRE](#)].
- [101] H1 collaboration, *Jet production in ep collisions at low Q^2 and determination of α_s* , *Eur. Phys. J. C* **67** (2010) 1 [[arXiv:0911.5678](#)] [[INSPIRE](#)].
- [102] H1 and ZEUS collaborations, *Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data*, *Eur. Phys. J. C* **75** (2015) 580 [[arXiv:1506.06042](#)] [[INSPIRE](#)].
- [103] J. Currie, T. Gehrmann and J. Niehues, *Precise QCD predictions for the production of dijet final states in deep inelastic scattering*, *Phys. Rev. Lett.* **117** (2016) 042001 [[arXiv:1606.03991](#)] [[INSPIRE](#)].
- [104] J. Currie, T. Gehrmann, A. Huss and J. Niehues, *NNLO QCD corrections to jet production in deep inelastic scattering*, *JHEP* **07** (2017) 018 [*Erratum ibid.* **12** (2020) 042] [[arXiv:1703.05977](#)] [[INSPIRE](#)].
- [105] NNPDF collaboration, *Parton distributions with theory uncertainties: general formalism and first phenomenological studies*, *Eur. Phys. J. C* **79** (2019) 931 [[arXiv:1906.10698](#)] [[INSPIRE](#)].
- [106] NNPDF collaboration, *Determination of the theory uncertainties from missing higher orders on NNLO parton distributions with percent accuracy*, *Eur. Phys. J. C* **84** (2024) 517 [[arXiv:2401.10319](#)] [[INSPIRE](#)].
- [107] NNPDF collaboration, *Fitting parton distribution data with multiplicative normalization uncertainties*, *JHEP* **05** (2010) 075 [[arXiv:0912.2276](#)] [[INSPIRE](#)].

- [108] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012 [[hep-ph/0201195](#)] [[INSPIRE](#)].
- [109] Z. Kassabov, E.R. Nocera and M. Wilson, *Regularising experimental correlations in LHC data: theory and application to a global analysis of parton distributions*, *Eur. Phys. J. C* **82** (2022) 956 [[arXiv:2207.00690](#)] [[INSPIRE](#)].
- [110] ATLAS collaboration, *Measurement of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **09** (2017) 020 [[arXiv:1706.03192](#)] [[INSPIRE](#)].
- [111] L.A. Harland-Lang, A.D. Martin and R.S. Thorne, *The impact of LHC jet data on the MMHT PDF fit at NNLO*, *Eur. Phys. J. C* **78** (2018) 248 [[arXiv:1711.05757](#)] [[INSPIRE](#)].
- [112] NNPDF collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [[arXiv:1706.00428](#)] [[INSPIRE](#)].
- [113] M.A.G. Aivazis, J.C. Collins, F.I. Olness and W.-K. Tung, *Leptoproduction of heavy quarks. 2. A unified QCD formulation of charged and neutral current processes from fixed target to collider energies*, *Phys. Rev. D* **50** (1994) 3102 [[hep-ph/9312319](#)] [[INSPIRE](#)].
- [114] M.A.G. Aivazis, F.I. Olness and W.-K. Tung, *Leptoproduction of heavy quarks. 1. General formalism and kinematics of charged current and neutral current production processes*, *Phys. Rev. D* **50** (1994) 3085 [[hep-ph/9312318](#)] [[INSPIRE](#)].
- [115] M. Krämer, F.I. Olness and D.E. Soper, *Treatment of heavy quarks in deeply inelastic scattering*, *Phys. Rev. D* **62** (2000) 096007 [[hep-ph/0003035](#)] [[INSPIRE](#)].
- [116] W.-K. Tung, S. Kretzer and C. Schmidt, *Open heavy flavor production in QCD: conceptual framework and implementation issues*, *J. Phys. G* **28** (2002) 983 [[hep-ph/0110247](#)] [[INSPIRE](#)].
- [117] ATLAS collaboration, *Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 367 [[arXiv:1612.03016](#)] [[INSPIRE](#)].
- [118] R.S. Thorne, *Effect of changes of variable flavor number scheme on parton distribution functions and predicted cross sections*, *Phys. Rev. D* **86** (2012) 074017 [[arXiv:1201.6180](#)] [[INSPIRE](#)].
- [119] S. Forte, E. Laenen, P. Nason and J. Rojo, *Heavy quarks in deep-inelastic scattering*, *Nucl. Phys. B* **834** (2010) 116 [[arXiv:1001.2312](#)] [[INSPIRE](#)].
- [120] S. Carrazza, J.M. Cruz-Martinez and R. Stegeman, *A data-based parametrization of parton distribution functions*, *Eur. Phys. J. C* **82** (2022) 163 [[arXiv:2111.02954](#)] [[INSPIRE](#)].
- [121] L. Del Debbio, T. Giani and M. Wilson, *Bayesian approach to inverse problems: an application to NNPDF closure testing*, *Eur. Phys. J. C* **82** (2022) 330 [[arXiv:2111.05787](#)] [[INSPIRE](#)].
- [122] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [[arXiv:1506.07443](#)] [[INSPIRE](#)].
- [123] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204 [[arXiv:1412.3989](#)] [[INSPIRE](#)].
- [124] R.D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [[arXiv:1207.1303](#)] [[INSPIRE](#)].
- [125] T. Sharma, *Exact NNLO corrections vs K-factors in PDF fits*, *PoS DIS2024* (2025) 043 [[arXiv:2408.10424](#)] [[INSPIRE](#)].

- [126] M.L. Mangano et al., *Physics at a 100 TeV pp collider: Standard Model processes*, [arXiv:1607.01831](#) [INSPIRE].
- [127] T. Cridge, L.A. Harland-Lang, A.D. Martin and R.S. Thorne, *QED parton distribution functions in the MSHT20 fit*, *Eur. Phys. J. C* **82** (2022) 90 [[arXiv:2111.05357](#)] [INSPIRE].
- [128] CTEQ-TEA collaboration, *The photon content of the neutron*, *JHEP* **04** (2024) 022 [[arXiv:2305.10497](#)] [INSPIRE].
- [129] NNPDF collaboration, *Photons in the proton: implications for the LHC*, *Eur. Phys. J. C* **84** (2024) 540 [[arXiv:2401.08749](#)] [INSPIRE].
- [130] NNPDF collaboration, *The path to N^3 LO parton distributions*, *Eur. Phys. J. C* **84** (2024) 659 [[arXiv:2402.18635](#)] [INSPIRE].
- [131] D0 collaboration, *Measurement of the electron charge asymmetry in $p\bar{p} \rightarrow W + X \rightarrow e\nu + X$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. D* **91** (2015) 032007 [Erratum *ibid.* **91** (2015) 079901] [[arXiv:1412.2862](#)] [INSPIRE].
- [132] ATLAS collaboration, *Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 367 [[arXiv:1612.03016](#)] [INSPIRE].
- [133] LHCb collaboration, *Measurement of the cross-section for $Z \rightarrow e^+e^-$ production in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **02** (2013) 106 [[arXiv:1212.4620](#)] [INSPIRE].
- [134] LHCb collaboration, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **08** (2015) 039 [[arXiv:1505.07024](#)] [INSPIRE].
- [135] LHCb collaboration, *Measurement of forward $Z \rightarrow e^+e^-$ production at $\sqrt{s} = 8$ TeV*, *JHEP* **05** (2015) 109 [[arXiv:1503.00963](#)] [INSPIRE].